Isolating and Assaying Unspecific Peroxygenase and Flavin Binding Enzymes for *in vitro* Terpenoid Biosynthesis

A THESIS SUBMITTED TO THE FACULTY OF THE UNIVERSITY OF MINNESOTA BY

Benjamin C. Hanson

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

Dr. Claudia Schmidt-Dannert, Advisor

May 2018

Acknowledgements

I would like to thank my family for all of their love and support. I am also grateful for the guidance of my mentor, Claudia, all of my amazing labmates, my committee members Romas Kazlauskas and Mark Distefano, and all of my friends. I am also grateful for the funding provided by the NIH Biotechnology Training Grant.

Abstract

Terpenoids are an exceptionally large family of natural products, and contain numerous bioactive members that are pharmaceutically important. While most research into terpenoids and their metabolism has thus far occurred in non-fungal organisms, chiefly plants, Basidiomycota (mushroom forming fungi) are well known as prolific producers of bioactive sesquiterpenoids, such as the potent anticancer compounds Illudin M and S. While natural products have traditional been a huge driver of pharmaceutical discovery, this natural abundance is often hampered by very low expression in the native host and slow growth or rarity of the host itself. This drives up financial and environmental costs, and in many cases makes the production of otherwise useful natural products impracticable.

One potential solution to this quandary is heterologous production of secondary metabolites in non-native hosts such as *Escherichia coli* and *Saccharomyces cerevisiae.* Another avenue is *in vitro* biocatalysis, wherein the biosynthetic enzymes of the native host are heterologously expressed, isolated, and used to perform synthesis outside of the cell. This approach avoids the fragility of the *in vivo* system and would allow the creation of combinatorial enzymatic pathways to create novel bioactive structures. With regards to constructing a sesquiterpenoid biosynthetic pathway, many stable terpene synthases have been isolated and shown to be active *in vitro*. However, the most well studied terpene scaffold modifying enzymes, cytochrome P450s, are notoriously difficult to heterologously express in active form. In order to obtain scaffold modifying oxygenases capable of being part of an *in vitro* terpenoid biosynthetic pathway, fungal oxygenases aside from cytochrome P450s were investigated. Specifically, unspecific peroxygenase

from *Agrocybe aegerita* and flavin binding oxidoreductases from Δ6-protoilludene biosynthetic gene clusters were expressed and assayed against the sesquiterpene scaffold.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

For millennia, humans have used plants and fungi as our primary source of salves and medicines. The bioactive chemical compounds of these organisms are termed natural products, also known as secondary metabolites. Unlike primary metabolites, which are essential for an organism's life and ability to function, secondary metabolites are not required for survival. Many secondary metabolites are chemical warfare agents which confer an evolutionary advantage against competitors, pathogens, and predators, and have cytotoxic properties which can be applied against bacteria, fungi, and cancer cells.[1] In more recent times, natural products became the basis for the majority of modern medicines. Aspirin, derived from salicylic acid isolated from willow bark,[2] morphine from opium poppy, quinine from *Cinchona succirubra*,[1] and penicillin from the mold *Penicillium notatum*,[3] are only some of a multitude of modern pharmaceuticals derived or isolated from natural sources. In the past 30 years, 61% of approved anticancer compounds and 49% of approved anti-infectives were derived from or inspired by natural products.[4]

Terpenoids, which are formed from five-carbon isoprenyl diphosphate molecules,[5] are considered to be the largest and most diverse class of natural products.[6] This class contains many pharmaceutically important compounds such as artemisinin (the basis of numerous anti-malarial drugs), paclitaxel (chemotherapy medication) and pleuromutilin (the source of the anti-biotic semi-synthetic derivatives tiamulin, valnemulin, and retapamulin).[7] Traditionally, terpenoids and other natural products are isolated by extraction from native host material. However, this harvesting method has significant drawbacks, chiefly that the supply of host material is often limited and

extraction yields are often very low.[8] This increases drug price and can also lead to environmental degradation. An example of this is the terpenoid paclitaxel, a very important anti-cancer drug, which was initially harvested from the bark of the Pacific Yew (*Taxus brevifolia*). Paclitaxel exists in low concentrations (0.01% - 0.05%) in the bark of *T. brevifolia.* In order to extract 1 kg of paclitaxel (capable of treating five hundred patients), up to 300 *T. brevifolia* trees must be killed to provide 10 tons of bark. As *T. brevifolia* grows slowly and is relatively uncommon, the strategy of direct extraction from bark is unsustainable both financially and ecologically.[9]

While total chemical synthesis of natural products is feasible for simpler molecules such as aspirin,[2] many natural products and especially terpenoids such as paclitaxel are very complex, making total synthesis impractical due to loss of yield over many steps and the production of inactive or toxic isomers.[8,10] If extraction from the native source material and total chemical synthesis have many drawbacks, a potentially viable alternate method is production of the target molecule in a heterologous host. This can be done *in vivo* by using recombinant DNA technology to express the enzymes of the natural product pathway in the heterologous host, which is typically *S. cerevisae* or *E. coli* as these organisms have been well characterized. Heterologous production of the artemisinin precursor artemisinic acid was achieved through the recombinant introduction of the nine genes of the mevalonate pathway, expression of the Artemisia annua amorphadiene synthase, a cytochrome P450 and its reductase, another cytochrome enzyme, and two dehydrogenases in *S. cerevisiae.* This resulted in the production of 25 g/L of artemisinic acid, which is extracted and modified by organic synthesis to artemisinin, a process currently undergoing large scale industrial implementation.[2]

In addition to *in vivo* production of natural products, these molecules can also be biosynthesized *in vitro* by heterologously expressing, purifying, and isolating the requisite enzymes and performing the reactions in an otherwise artificial system. Some benefits of the *in vitro* approach include the ability to use substrates and produce products that may be toxic to an *in vivo* host, and the possibility of utilizing reaction media that is incompatible with an *in vivo* system (such as organic solvents). *In vitro* biocatalysis also allows the order of enzymatic reactions to be easily rearranged, and non-native enzymes can be used to further modify the natural product, or enzymes from different pathways can be utilized in tandem to create molecules with novel structures and bioactivities. However, *in vitro* biocatalysis is challenged by the difficulty of isolating enzymes that are active and stable *in vitro*, and the need to externally supply expensive co-factors.[8]

Heterologous production of terpenoids, *in vivo* or *in vitro*, could produce a much needed boost to pharmaceutical research and production. One important but relatively untapped source of terpenoid natural products is Basidiomycota, or mushroom forming fungi.[11] Basidiomycota have played a crucial role in traditional medicine since ancient times, they are known to produce a great range of natural products with antibacterial, anti-cancer, and anti-fungal activity, and terpenoids are one of the most common classes they produce. [5,11,12,13] Most investigation into fungal terpenoid metabolism has focused on Ascomycota (filamentous fungi), with relatively little research into the metabolism of Basidiomycota natural products. This is likely due to the fact that Ascomycota are easy to grow in a laboratory and are genetically tractable, while Basidiomycota are difficult or impossible to grow in the laboratory and with a few exceptions are not genetically tractable. However, with advances in fungal genome sequencing and synthetic biology,

elucidation of Basidiomycota natural product metabolism is becoming more and more feasible.[5]

If heterologous production of basidiomycete terpenoids is to be achieved, their metabolic pathways must first be characterized and then transferred to the production host. While specific enzymes may still need to be identified, the general outline of terpenoid biosynthesis is known: (I): synthesis of isoprenyl diphosphate precursors, (II), cyclization of the precursor molecule into the hydrocarbon terpene backbone, (III) scaffold modification. In fungi the five carbon precursor isopentenyl diphosphate (IPP) is synthesized from acetyl-CoA through the mevalonate pathway, and a portion of the IPP produced is isomerized to dimethylallyl pyrophosphate (DMAPP) by IPP isomerase.[5,14] An alternative route to IPP and DMAPP, known as the 1-deoxy-D-xylulose 5-phosphate (DXP) pathway, is used by *E. coli* and some other bacteria, while plants utilize both pathways.[2] IPP units are sequentially added to DMAPP in a 1'-4 condensation reaction catalyzed by isoprenyl diphosphate synthase, resulting in geranyl pyrophosphate (GPP), farnesyl pyrophosphate (FPP), or geranylgeranyl pyrophosphate (GGPP), which contain 10, 15, and 20 carbons respectively.[5]

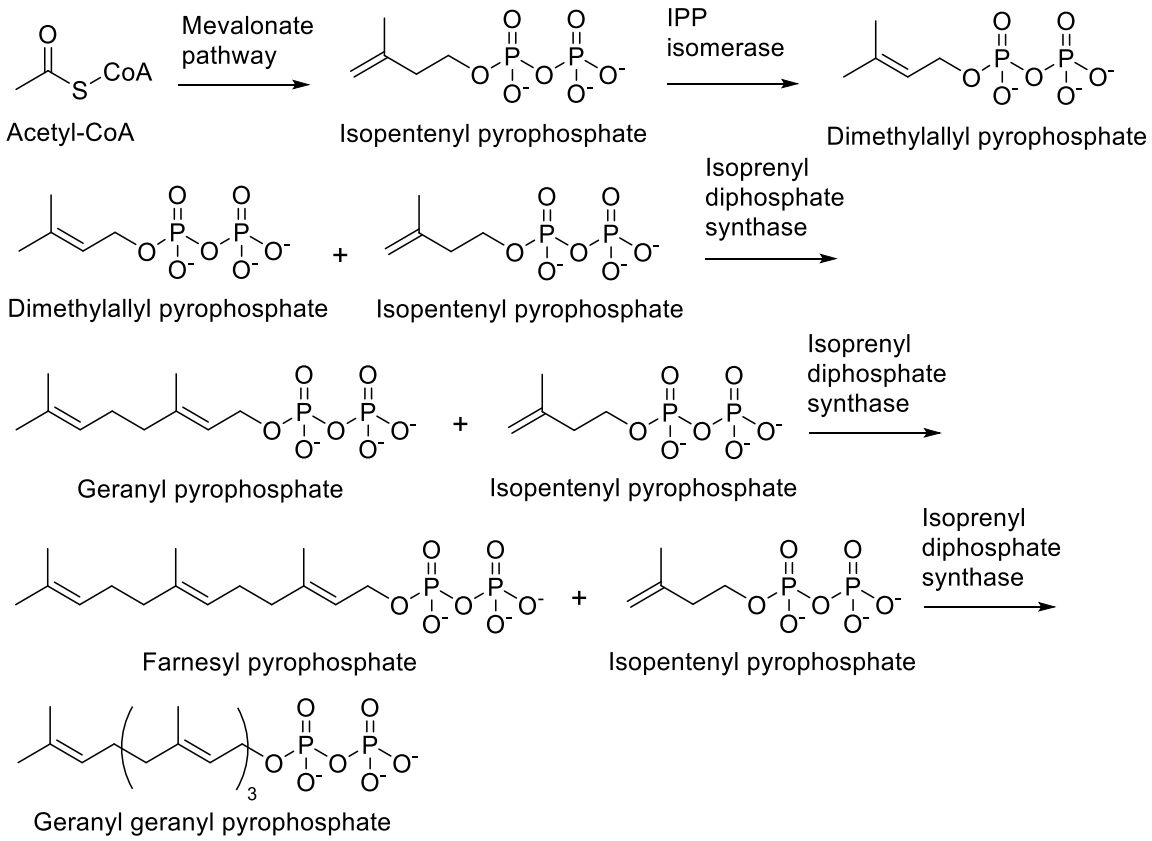


Figure 1.1 - Biosynthesis of isoprenyl diphosphate terpenoid precursors

These precursors are dephosphorylated by terpene synthases (also known as terpene cyclases), and undergo an ionization dependent (Class I) or protonation dependent (Class II) cyclization cascade and yield the hydrocarbon (10 carbon) mono-, (15 carbon) sesqui-, (20 carbon) di-, or (30 carbon) triterpene scaffolds.[5,13] The terpene scaffolds are then acted upon by scaffold modifying enzymes such as cytochrome P450 dependent monooxygenases (CYPs), oxidoreductases and transferases.[13] Basidiomycota are most well known for producing biologically active sesquiterpenoids,[7] which contain 15 carbons and are synthesized from farnesyl pyrophosphate, typically by Class I terpene synthases. The sesquiterpene synthase mediates metal ion induced departure of pyrophosphate from FPP, resulting in a highly reactive farnesyl carbocation that

undergoes ring closure at the 1,6, 1,10, or 1,11 position followed by further cyclization reactions and ring rearrangements which form the sesquiterpene scaffold.[13]

While Basidiomycota do produce a number of sesquiterpenes derived from 1,6 and 1,10 cyclized cations, most of the medically relevant basidiomycete sesquiterpenoids are derived from the 1,11 cyclized trans-humulyl cation. Sesquiterpenoid classes derived from this cation include the caryophyllanes, africananes, tremulanes, humulanes, sterpuranes, hirsutanes, pentalenene derivatives, and Δ6-protoilludene derivatives (see Figure 2 below).[1]

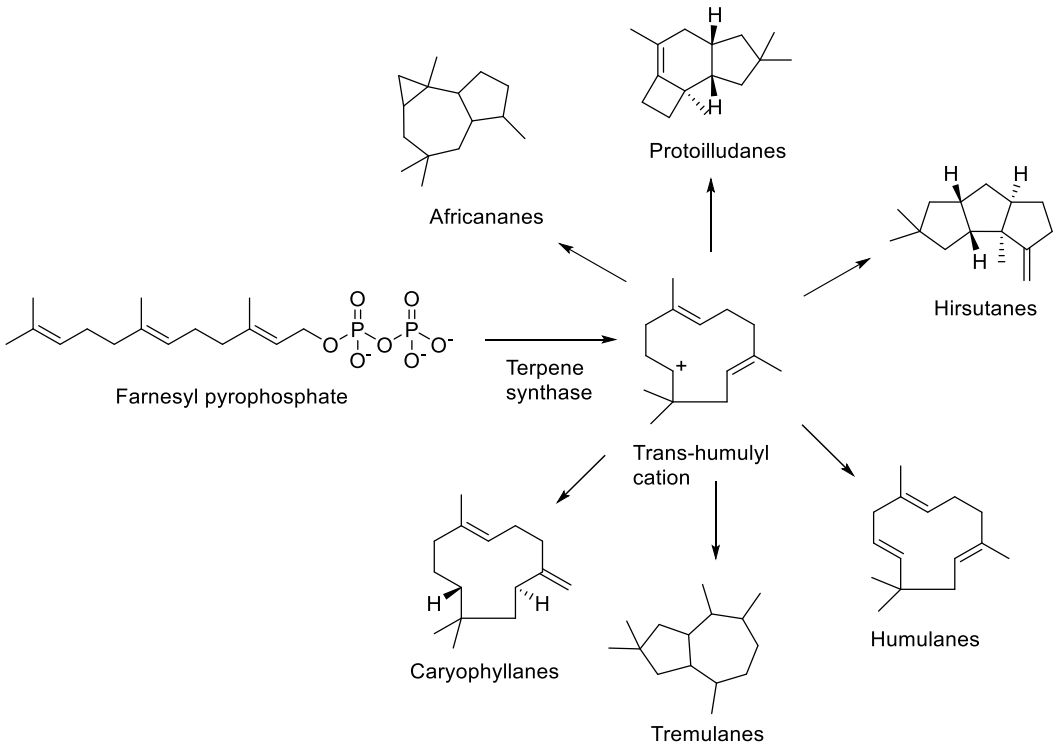


Figure 1.2 - Different classes of terpenoid scaffolds derived from the 1,11 cyclized trans-humulyl cation.

The Δ6-protoilludene scaffold in particular is the precursor for a variety of bioactive terpenoids, such as the illudanes, marasmanes, lactaranes, and fommanosanes (see Figure 1.3 below). Two of the most promising pharmaceutical candidates from

Basidiomycota, the *Omphalatus olearius* illudanes known as Illudin M and S, are currently being developed as anti-cancer therapeutics.[13]

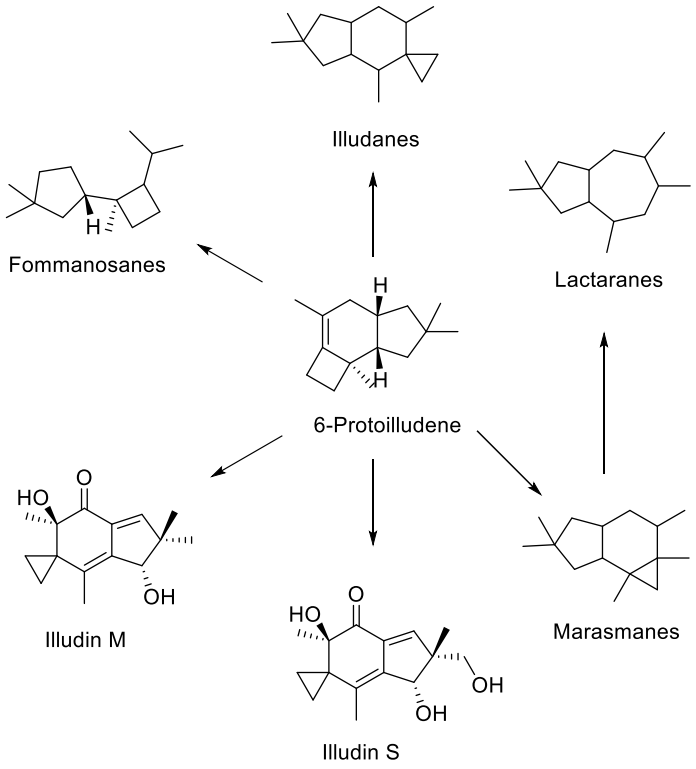


Figure 1.3 - Different bioactive terpenoid classes derived from the Δ6-protoilludene scaffold.

Achieving heterologous biosynthesis of fungal terpenoids would enable medical investigation of many bioactive compounds that have been isolated but are only produced in minute quantities, and in the case of prospective pharmaceuticals such as Illudin M and S could provide the ability to produce future drugs on an industrial scale. If fungal terpenoid biosynthesis could be achieved *in vitro*, it would circumvent the inherent fragility of living systems that can complicate *in vivo* biosynthesis. In addition, with an *in vitro* biosynthesis pathway the biocatalytic enzymes could be rearranged, utilized in another pathway, or new enzymes could be added to the pathway to generate a variety of novel molecules with potential bioactivities. In this thesis, I will give a description of my

attempts to create a combinatorial, enzymatic pathway for the *in vitro* biocatalysis of fungal sesquiterpenoids.

The sesquiterpene precursor FPP is commercially available and can be added directly to the biocatalytic reaction mixture. In order to achieve an enzymatic cascade, sesquiterpene synthases and scaffold modifying enzymes that are active and stable *in vitro* are required. A number of fungal sesquiterpene synthases have been heterologously expressed and shown to be both stable and active *in vitro*. The Schmidt-Dannert lab has previously expressed and isolated a wide variety of sesquiterpene synthases from the mushrooms *Coprinus cinereus,*[15] *Omphalotus olearius*[16] and *Stereum hirsutum*[17] including a number of Δ6-protoilludene synthases which are available for use in the lab. A number of these sesquiterpene synthases, such as the α-cuprenene synthase Cop6 from *C. cinereus*, the Δ6-protoilludene synthases Omp6 and Omp7 from *O. olearius*, and the prototilludene synthases Stehi1|25180, Stehi1|64702 and Stehi1|73029, from *S. hirsutum* are located in large biosynthetic gene clusters containing a number of scaffold modifying enzymes.[15,16,17]

CYPs are the most common scaffold modifying enzyme in sesquiterpenoid biosynthesis,[7] and indeed the sesquiterpene synthase gene clusters from *C. cinereus, O. olearius,* and *S. hirsutum* contain a number of CYP genes.[15,16,17] However, the use of fungal CYPs in heterologous enzymatic pathways is complicated by the difficulty of expressing active CYPs. Most eukaryotic CYPs are membrane bound[18] and when expressed in heterologous hosts these enzymes often suffer from low expression levels, instability, protein misfolding, and aggregation into inclusion bodies. Thus in heterologous expression, it is common for fungal CYPs to not express at all or express in

inactive form.[58] While two *C. cinereus* cytochrome P450 enzymes (Cox1 and Cox2) in the biosynthetic cluster of Cop6 appeared to have activity on α-cuprenene when coexpressed with Cop6,[15] other attempts to heterologously express CYPs from the sesquiterpenoid biosynthetic gene clusters of *C. cinereus, O. olearius,* and *S. hirsutum* have not yielded activity (unpublished data).

As Δ6-protoilludene is the precursor to many of the most bioactive fungal sesquiterpenoids, I focused my efforts on creating a biosynthetic pathway to Δ6-protoilludene derivatives. In addition to the potential pharmaceutical benefits of synthesizing Δ6-protoilludene derivatives, it would also be interesting to discover an enzyme which causes the opening and rearrangement of Δ6-protoilludene's strained cyclobutane ring. This is a crucial yet mysterious step that results in much of the structural diversity of Δ6-protoilludene derivatives. Δ6-protoilludene is easily supplied *in vitro* by the activity of the Stehi7 Δ6-protoilludene synthase from *S. hirsutum.*. To produce the final terpenoid product, it is necessary to assemble an ensemble of scaffold modifying enzymes with different activities. As CYP enzymes have proven to be difficult to express in active form, alternate fungal scaffold modifying enzymes were heterologously expressed and investigated for activity against sesquiterpenes. The remainder of this thesis describes an investigation of the activity of A) a mutant of the unspecific peroxygenase UPO from the mushroom *Agrocybe aegerita,* which was evolved by another lab group for expression in *S. cerevisiae*[19] and B) flavin-binding enzymes found in the biosynthetic gene clusters of the *O. olearius* Omp7 Δ6-protoilludene synthase and the *S. hirsutum* Stehi7 Δ6-protoilludene synthase.

Chapter 2

Expression and investigation of *Aae*UPO mutant PaDa I

Introduction

A. aegerita unspecific peroxygenase (UPO) is a mono-peroxygenase that I investigated for activity against the Δ6-protoilludene sesquiterpene scaffold. Broadly, peroxygenases are enzymes that transfer a peroxide borne oxygen atom to substrates.[20] UPO was first isolated from *Agrocybe aegerita*, a popular edible basidiomycete which grows on wood and bark mulch and is found throughout the Mediterranean region.[21] Interest in UPO increased greatly as it was shown to be capable of oxygenating linear, branched and cyclic alkenes and alkanes (with alkanes ranging in size from propane (C_3) to hexadecane (C_{16})),[22,23] aromatic compounds such as naphthalene,[24] and benzene,[25] heterocycles such as dibenzofuran,[26] and ethers (causing cleavage)[27] among other substrates. UPOs are capable of performing dealkylation, hydroxylation, epoxidation, aromatization, sulfoxidation, dechlorination, and halide oxidation.[20] The selective oxygenation of poorly activated C – H bonds is considered a "dream reaction" in organic chemistry, as it is difficult to accomplish and very desired within industrial and pharmaceutical synthesis.[28] The other principal group of enzymes capable of functionalizing C – H bonds are cytochrome P450s, which have been well studied. Often, the products of unspecific peroxygenase reactions are similar to human cytochrome P450s.[20] However, unlike CYPs which are membrane bound, fragile, and co-factor dependent, UPOs are soluble, excreted extracellularly, stable, and only require low (1-2 mM) amounts of H_2O_2 in order to be active.[19,29] In addition, use of UPO can enable

increased (when compared with traditional metal catalysts) or near complete stereoselectivity or regioselectivity.[23,30]

UPOs are one of two types of enzymes which are part of the heme thiolate peroxidase superfamily, the other enzyme type being the chloroperoxidase.[31] Chloroperoxidase is secreted by the filamentous fungus *Caldariomyces fumago*, and catalyzes oxidative chlorination,[32] as well as epoxidation of linear alkenes and hydroxylation of benzylic carbons. However, it is unable to oxygenate stronger C – H bonds such as those found in aromatic compounds or alkanes.[20] Heme thiolate peroxidases all contain a heme domain in the active site, and a thiolate group which acts as a ligand to the heme Fe^{III} ion.[19]

After the discovery of *A. aegerita* UPO (*Aae*UPO), other UPOs were also isolated from the fungi *Coprinus radians* (*Cra*UPO)[33] and *Marasmius rotula* (*Mro*UPO).[34] In addition, it is known that at least eight other mushrooms secrete UPOs. When genetic databases are searched for UPO like sequences, approximately 2000 putative UPO sequences are found in fungi.[20] These UPOs can be divided into two groups based on their length: short UPOs are on average 29 kDa and contain CPO and *Mro*UPO, while long UPOs are on average 44 kDa and contain *Aae*UPO and *Cra*UPO.[20]

UPOs transfer oxygen to substrates in a similar manner as CYPs (the "peroxide shunt" pathway), and also oxidize substrates in a similar manner as heme peroxidases. In this way, UPOs potentially represent an evolutionary "missing link" between the CYP and heme peroxidase enzyme classes[20] (see Figure 4 below). In the UPO oxidation mechanism, an oxoiron(IV) protoporphyrin radical cation intermediate is the species that reacts with the substrate.[35,36]

Balanced equation for UPO reaction:[20] R-H + H_2O_2 ⇌ R-OH + H_2O

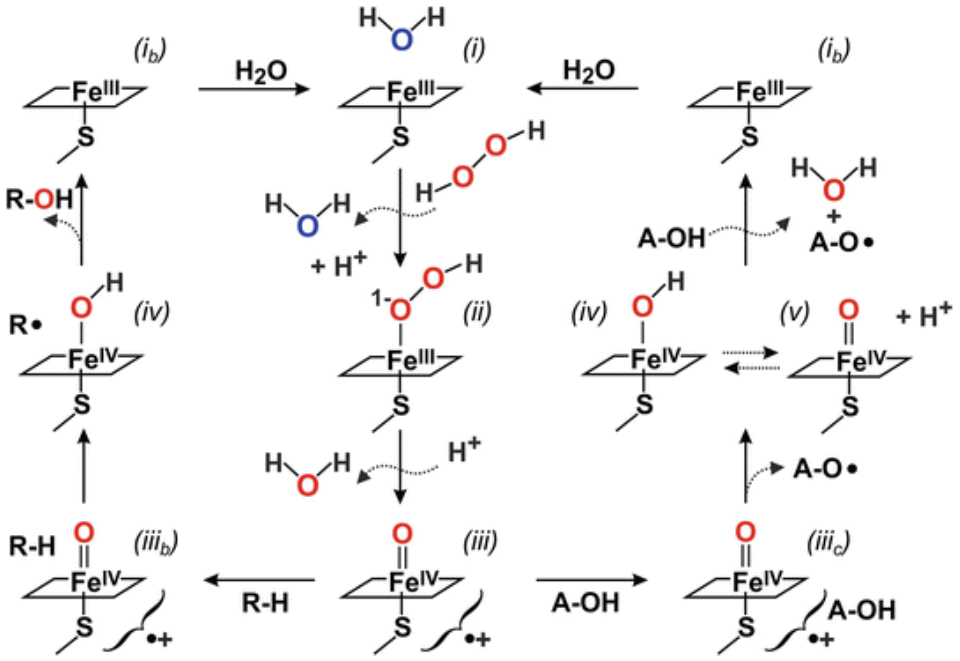


Figure 2.1 – UPO reaction mechanism. Oxidation reaction mechanism for unspecific peroxygenase, oxygen transfer reaction shown on left and oxidation without oxygen transfer shown on right (taken from Hofrichter et al., 2015).[20]

UPOs including *Aae*UPO were originally harvested directly from fungal culture. In order to produce AaeUPO in a manner more amenable to industrial adoption, and in order to enable expression of *Aae*UPO in a system that would allow the creation of mutant *Aae*UPO, Molina-Espeja et al., used directed evolution to optimize *Aae*UPO for secretion in *S. cerevisiae*.[19] Mutagenic PCR was used over five generations to produce a mutant AaeUPO (termed PaDa I), with four mutations in the signal peptide and five in the body of the enzyme. This process resulted in an increase of UPO expression by 1,114 fold and a specific activity increase of 3.6 fold, and a total secretion level of 7.8 mg/L. Glycosylation increased from 22% for the wild type to 30% for the PaDa I mutant, and the mutant was both active and highly stable in the presence of organic cosolvents.[19] The heavy glycosylation increases UPOs stability, which is one reason it was expressed in

yeast (which can glycosylate proteins, while *E. coli* cannot). Peroxygenase activity was measured by assaying the PaDa I – UPO mutant against 5-nitro-1,3-benzodioxole (NBD), transforming NBD into yellow colored 4-nitrocatechol, the increase of which can be measured spectrophotometrically at 425 nm.[37] Peroxidase activity was measured by assaying PaDa I – UPO against 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid (ABTS), which forms a stable green radical form of ABTS whose formation can be measured spectrophotometrically at 418 nm.[38] Due to the higher sensitivity of the ABTS assay and the fact that peroxidase and monooxygenase activities of PaDa I-UPO are closely linked, Molina-Espeja et al. used ABTS activity to measure total enzyme activity.[19] As *Aae*UPO is capable of oxyfunctionalizing a wide variety of hydrocarbons of varying sizes, including alkanes and alkenes, I hypothesize it will be able to modify sesquiterpene scaffolds such as Δ6-protoilludene. Previously, AaeUPO was shown to have activity against the monoterpene limonene, producing epoxylimonene and carveol, the same products produced by liver CYPs.[23] In order to obtain UPO to perform sesquiterpene activity assays with, I synthesized the gene coding for the PaDa I – UPO mutant, transferred it to a pESC-ura vector, and expressed it in *S. cerevisiae.*

Materials and Methods

Cloning of PaDa I – UPO gene into pESC-ura expression vector

The PaDa I – UPO mutant gene was designed according to the description in the original paper[19] and synthesized through Invitrogen's GeneArt® service. The gene was PCR amplified from the pMX plasmid it arrived in, and cloned with yeast recombinational cloning (utilizing BamHI and HindIII restriction enzymes) into a pESC-

ura vector (from Stratagene) under the control of the Gal1 (galactose inducible) promoter. All primers (see Table 2.2) were ordered from Integrated DNA Techologies, and Phusion® polymerase and restriction enzymes were purchased from New England Biolabs (NEB). Thermo Fisher Scientific Top10 cells were used for plasmid production during cloning. GoGreen® Taq polymerase from Promega was used to screen colonies for the presence of desired genes, and sequencing was performed at the University of Minnesota Genomics Center. The PaDa I - UPO containing pESC-ura plasmid was transformed into a BJ5465 protease deficient yeast strain (purchased from American Type Culture Collection, i.e. ATCC) using an in house transformation protocol which is described in the Supplemental Materials section.

Table 2.1 – Plasmids and strains used in Chapter 2

Plasmid name	Gene and promoter information	Selectable marker	Source
PMX_PaDa_I	Contains PaDa I – UPO gene	Kanamycin resistance	Synthesized by GeneArt® (Invitrogen), PaDa I –UPO gene from Molina-espeja et al[19]
pESC-ura	Contains the Gal1 and Gal 10 promoters	Ampicillin resistance, URA3* marker	From Schmidt-Dannert laboratory collection, originally from Stratagene
pESC-ura_PaDa_I	Contains PaDa I – UPO gene under control of the Gal1 promoter	Ampicillin resistance, URA3	This study
pESC-ura_PaDaI_A1	Contains PaDa I – UPO gene with Kozak sequence A1 (see Table 2.6) under control of the Gal1 promoter	Ampicillin resistance, URA3	This study
pESC-ura_PaDaI_A2	Contains PaDa I – UPO gene with Kozak sequence A2 (see Table 2.6) under control of the Gal1 promoter	Ampicillin resistance, URA3	This study
pESC-ura_PaDaI_A3	Contains PaDa I – UPO gene with Kozak sequence A3 (see Table 2.6) under control of the Gal1 promoter	Ampicillin resistance, URA3	This study
pESC-ura_PaDaI_A4	Contains PaDa I – UPO gene with Kozak sequence A4 (see Table 2.6) under control of the Gal1 promoter	Ampicillin resistance, URA3	This study
pESC-ura_PaDaI_A5	Contains PaDa I – UPO gene with Kozak sequence A5 (see Table 2.6) under control of the Gal1 promoter	Ampicillin resistance, URA3	This study
pESC-ura_PaDaI_A6	Contains PaDa I – UPO gene with Kozak sequence A6 (see Table 2.6) under control of the Gal1 promoter	Ampicillin resistance, URA3	This study
Strain	Description		Source
BJ5465	Protease deficient *S. cerevisiae* strain		ATCC
Top10	Chemically competent *E. coli* used for plasmid production		ThermoFisher Scientific

*URA3 allows growth on minimal media without added uracil.

Table 2.2 – Primers used in Chapter 2

Primer name	5' – 3' Sequence	Function
PaDa I F *	TATACCTCTATACTTTAACGTCAAGGAGAAAAAACCCCG	Amplifies PaDa I-UPO gene from PMX_PaDa_I plasmid, amplicon cut with BamHI and HindIII restriction enzymes, used to clone PaDa I – UPO into pESC-ura to produce pESC-ura_PaDa_I plasmid.
PaDa I R	GGTTAGAGCGGATCTTAGCTAGCCGCGGTACCAAGCTTACTCG	
A1 F	TGGAATATTTTCCCCTGTTCC	Performs site directed mutagenesis on pESC-ura_PaDa_I plasmid to produce pESC-ura_PaDaI_A1 plasmid.
A1 R	TGGTTGAGTCGTATTACGGATC	
A2 F	ATACGACTCAACCATGAAATATTTTCCCCTGTTC	Performs site directed mutagenesis on pESC-ura_PaDa_I plasmid to produce pESC-ura_PaDaI_A2 plasmid.
A2 R	TACGGATCCGGGGTTTTT	
A3 F	ATGTCTTATTTTCCCCTGTTCCCAAC	Performs site directed mutagenesis on pESC-ura_PaDa_I plasmid to produce pESC-ura_PaDaI_A3 plasmid.
A3 R	TTTTTTGTCGTATTACGGATCCGG	
A4 F	AATGAAATATTTTCCCCTGTTCCC	Performs site directed mutagenesis on pESC-ura_PaDa_I plasmid to produce pESC-ura_PaDaI_A4 plasmid.
A4 R	TTTTTGTCGTATTACGGATCCGG	
A5 F	GTCTTATTTTCCCCTGTTCCCAAC	Performs site directed mutagenesis on pESC-ura_PaDa_I plasmid to produce pESC-ura_PaDaI_A5 plasmid.
A5 R	ATTATTGAGTCGTATTACGGATC	
A6 F	ATACGACTCAATAATGAAATATTTTCCC	Performs site directed mutagenesis on pESC-ura_PaDa_I plasmid to produce pESC-ura_PaDaI_A6 plasmid.
A6 R	TACGGATCCGGGGTTTTT	

* = F indicates forward primer, R indicates reverse primer

S. cerevisiae media

Filter sterilized minimal media for yeast expression contained 6.7 g of yeast nitrogen base, 1.92 g of yeast synthetic drop-out medium supplement without uracil, 20 g raffinose, 100 μg/mL ampicillin, and ddH_2O to 1 L. Selective yeast plates contained 6.7 g of yeast nitrogen base, 1.92 g of yeast synthetic drop out medium supplement without uracil, 20 g autoclaved bacto agar, 20 g glucose, 100 μg/mL ampicillin and ddH2O to 1,000 mL. Sterile expression media contained 11 g of yeast extract and 22 g peptone in 720 mL ddH_2O, autoclaved separately as "YP media," 67 mL 1 M filtered KH_2PO_4 pH

6.0 buffer, 111 mL 20% filtered galactose, 22 mL filtered $MgSO_4$ 0.1 M, 31.6 mL absolute ethanol, 100 μg/mL ampicillin, and ddH2O to 1,000 mL.

Expression of PaDa I – UPO in *S. cerevisiae*

The yeast expression procedure was adapted from that described in the original paper that produced PaDa I – UPO.[19] A yeast colony expressing PaDa I – UPO or a mutant was picked from a selective plate and used to inoculate 20 mL minimal media cultures in 125 mL flasks. These cultures were incubated for 48 h at 30ºC and 220 RPM, and then used to inoculate a second set of 20 mL minimal media cultures at a starting OD600 of 0.2. Cells were grown for two doubling times (approximately 6-8 hours) and then 45 mL of sterile expression media was inoculated with 5 mL of culture. Expression cultures were grown for 72 h at 25ºC and 220 RPM (unless otherwise noted), and then harvested by centrifugation at 4,500 RPM and 4ºC for 10 minutes. The PaDa I – UPO containing supernatant was filter sterilized with a 0.2 micron filter. The protein expression level was determined by SDS-PAGE analysis, with overnight staining in Coomassie Biosafe® stain. PaDa I – UPO supernatant was occasionally concentrated by centrifugation using a Millipore Amicon® Ultra centrifugal filter with a 10,000 Da molecular weight cut off limit.

PaDa I – UPO activity assay

PaDa I – UPO's peroxygenase activity was measured by assaying against NBD (5-nitro-1,3-benzodioxole),[37] while its peroxidase activity was measured by an ABTS (2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) assay.[38] The ABTS and NBD reaction mixtures were prepared, 10 μL of PaDa I – UPO supernatant was placed in a

plastic cuvette, the reaction mixture was added, and spectrophotometric measurements were recorded at 418 nm (ABTS) or 425 nm (NBD) every 30 seconds for 2 minutes. ABTS reaction mixture contained 100 mM sodium phosphate/citrate buffer at pH 4.4, 0.3 mM ABTS and 2 mM H_2O_2, while NBD reaction mixture contained 100 mM potassium phosphate buffer pH 7.0, 1 mM NBD, 15% acetonitrile and 1 mM H_2O_2. Total activity of the supernatant was measured using ABTS activity, which was defined as the amount of enzyme that oxidizes 1 μmol of ABTS per min in 100 mM sodium phosphate/citrate buffer pH 4.4 containing 2 mM H_2O_2. This was calculated from the average ΔOD418/min of three measurements in the ABTS assay, according to the Beer Lambert law and the extinction coefficient of ABTS radical cation, which is 36,000 $M^{-1}cm^{-1}$.

Purification of PaDa I – UPO

In order to concentrate and partially purify PaDa I – UPO, the PaDa I – UPO supernatant was precipitated with ammonium sulfate. Amount of supernatant and ammonium sulfate used in each precipitation is indicated in the Results and Discussion section. Two precipitations were performed, an initial precipitation with a lower concentration of ammonium sulfate, after which the precipitate was discarded, and a final precipitation, after which the precipitate was kept and the supernatant was discarded. The precipitated PaDa I – UPO was resuspended in 10 mM sodium phosphate/citrate pH 4.3 buffer (buffer A), and either dialyzed or run through a PD 10 desalting column from Amersham Biosciences (according to the manufacturers protocol) to remove ammonium sulfate in preparation for cation exchange chromatography. The desalted PaDa I – UPO solution was filtered and loaded on to a strong cation-exchange column (5 mL HiTrap SP HP) pre-equilibrated with buffer A. The proteins were eluted with a gradient of 0 to 25 %

1 M NaCl, at a flow rate of 1 mL per minute, over 55 minutes, and then with a gradient of 25% to 100% NaCl at 1mL/minute over 5 minutes.

Alteration of Kozak sequences

Kozak sequences were altered by site directed mutagenesis with the Q5® kit from NEB (New England Biolabs). All primers were designed with the NEB software NEBaseChanger®, and ordered from Integrated DNA Techologies. Top10® *E. coli* cells from Thermo Fisher Scientific were used for cloning, GoGreen® Taq polymerase from Promega was used to screen colonies for the presence of desired genes, and sequencing was performed at the University of Minnesota Genomics Center.

GC/MS analysis of PaDa I activity against Δ6-protoilludene

PaDa I – Δ6-protoilludene assays were set up in glass GC/MS vials as follows:

10 mM Potassium Phosphate buffer pH 7.0: 564 µL
1.9 mg/mL Δ6-Protoilludene synthase: 10 µL
100 mM MgCl2 = 70 µL
PaDa I (or negative control) concentrated supernatant = 25 µL
FPP = 14 µL
H_2O_2= 1.75 µL

After shaking for 4 hours, the rubber septum of the GC/MS vial was pierced and a 100 µM polydimethylsiloxane (PDMS) fiber was inserted into the headspace (for volatile analysis) or the reaction solution and allowed to sample for 10 minutes. The PDMS fiber was then inserted into the port of a HP GC 7890A Gas chromatograph coupled to a mass spectrometer with a HP MSD triple axis detector.

PaDa I – limonene assays were set up in glass GC/MS vials as follows:

Acetone : 240 µL
Limonene : 0.23 µL

PaDa I (or negative control) concentrated supernatant : 15 µL
H_2O_2 : 1 µL
Fill to 400 µL with 10 mM Potassium phosphate pH 7.0

Reactions were shaken overnight, extracted with a half of the reaction volume's worth of hexane (1:2 extraction), and 1 µL of the hexane extract was deposited in the injection port by syringe. GC/MS programs were run in which the oven temperature began at 100 ˚C and reached 250 ˚C, for the volatile PaDa I – protoilludene sample and limonene sample this occurred over 15 minutes and for the PaDa I – protoilludene sample this occurred over 38 minutes.

Results and Discussion

Expression of *Agrocybe aegerita* unspecific peroxygenase mutant PaDa I

The synthesized gene coding for the *Agrocybe aegerita* unspecific peroxygenase mutant PaDa I (hereafter referred to as PaDa I-UPO) was inserted into a pESC-ura expression vector under control of the GAL1 (galactose inducible) promoter, creating vector pESC-ura_PaDa_I. pESC vectors have been widely and successfully used for expression in *S. cerevisiae*[39] and galactose inducible promoter based expression systems are among the strongest used in yeast.[40] This, and the expression of PaDa I-UPO under the control of a GAL1 promoter in the original paper[19] motivated the choice of expression vector.

With pESC-ura_PaDa_I, PaDa I-UPO was expressed in the protease deficient *S. cerevisiae* strain BJ5465. The culture supernatant was assayed with ABTS and NBD reagents to determine peroxidase and mono(per)oxygenase activity, respectively. Both assays returned positive results for the PaDa I-UPO expression culture supernatant, and negative (no color) results for the same culture prior to galactose induction and for

expression supernatant from a BJ5465 culture containing empty pESC-ura. Thus, it appears that control of PaDa I-UPO is tight and the negative control has no activity. As shown in Table 2.3 below, PaDa I-UPO containing supernatant from this initial expression had 17% of the ABTS activity reported in the original paper.[19]

Table 2.3 - Comparison of total activity of expression culture supernatant to reported activity. The number after the ± sign is the standard deviation.

Total activity (U/mL) of expression culture supernatant (first expression)	Total activity (U/mL) of PaDa I-UPO expression culture supernatant from Molina-Espeja, et al.[19]	Comparison of activity in expression culture supernatant to that reported in Molina-Espeja et al.[19] (%)
0.56 ± 0.045	3.4	17%

Despite the presence of UPO activity in the expression culture supernatant and lack of activity in the empty vector control supernatant, SDS-PAGE analysis revealed no distinct protein bands between the two supernatants. In addition, in contrast to reported SDS-PAGE gels of PaDa I-UPO,[19] a dark band was not observed in the supernatant at the expected weight of 51.1 kDa (Figure 2.2).

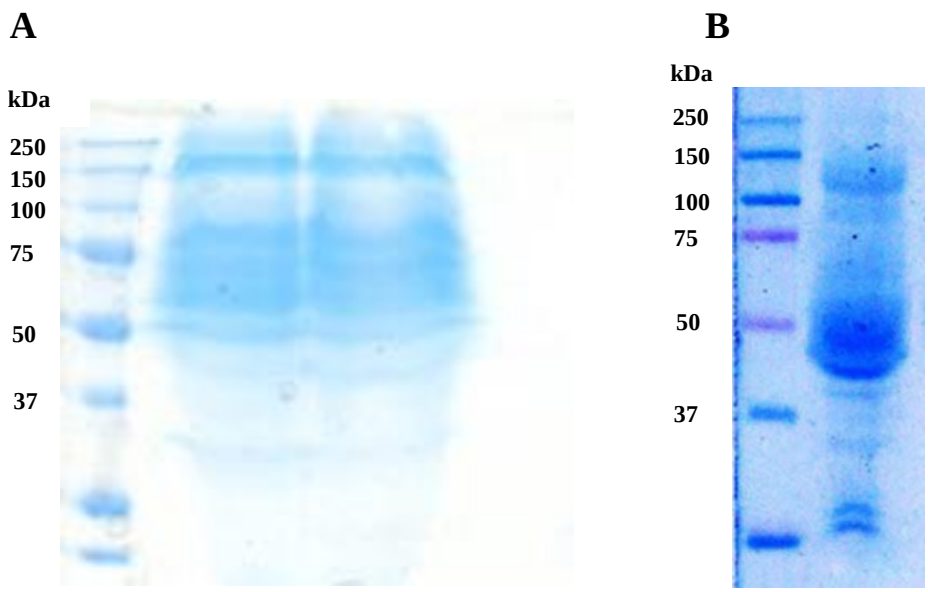


Figure 2.2 - SDS PAGE gels of first PaDa I-UPO expression A) this study (initial expression in BJ5465 protease deficient strain concentrated 72 x by filtration) and **B)** original study which produced PaDa I-UPO (from Molina-Espeja et al., 2014). In **B)** the band reported as PaDa I-UPO is located at approximately 50 kDa.

Both the lack of a distinct band for PaDa I-UPO and the low activity of the supernatant (compared to reported activity) indicated that PaDa I-UPO was being produced in insufficient quantity. In order to increase PaDa I – UPO production, expression temperature and time of expression induction were varied as these variables have been shown to affect protein expression yield.[41,42] One set of cultures was expressed at 25˚C (as in Molina-Espeja et al.)[19] while another was expressed at 20˚C to determine whether a lower expression temperature could improve protein solubility. Minimal media cultures were used to inoculate expression cultures (containing galactose as inductant) after two, two and a half, and three culture doubling times (Figure 2.3).

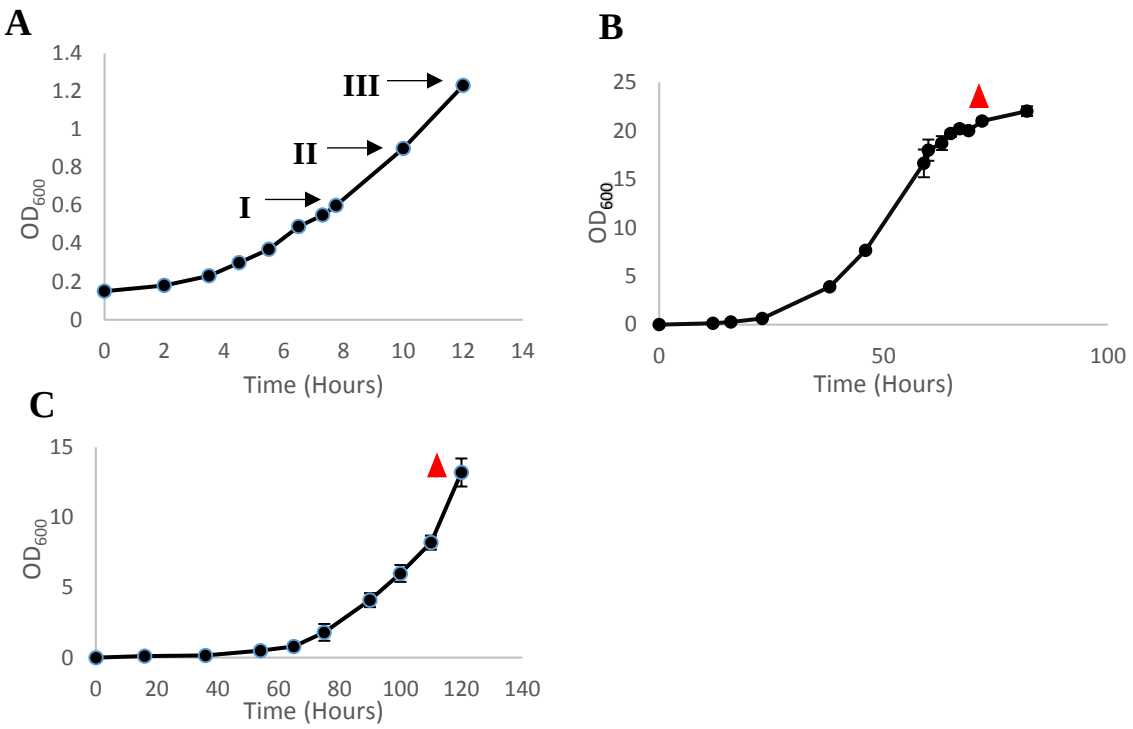


Figure 2.3 - Growth of *S. cerevisiae* BJ5465 cultures expressing PaDa I-UPO. A) Growth in minimal media of a representative culture. **I, II,** and **III** indicate induction at two, two and a half, and three doubling times respectively. **B)** Growth of expression culture at 25˚C and **C)** 20˚C, with a red triangle by a timepoint signifying time of supernatant harvest. Error bars represent standard deviation.

Following expression, cultures were tested for UPO activity with ABTS as the substrate (Table 2.4).

Table 2.4 - Total activity (measured through activity on ABTS) of PaDa I-UPO expressing BJ5465 *S. cerevisiae* cultures. The number of each culture indicates the initial yeast colony used for inoculation (three replicates performed in total). The Roman numeral subscript of each culture indicates the timepoint of induction, with I, II, and III indicating induction at two, two and a half, and three doubling times respectively. The number after the ± symbol is the standard deviation.

25˚C Culture	1_I	1_{II}	1_{III}	2_I	2_{II}	2_{III}	3_I	3_{II}	3_{III}
Total activity (U/mL)	1.43± 0.0753	1.28± 0.0588	1.04± 0.12	1.42± 0.043	1.12± 0.046	1.07± 0.049	1.18± 0.057	1.63± 0.058	0.938± 0.077
Culture at 20˚C	1_I	1_{II}	1_{III}	2_I	2_{II}	2_{III}	3_I	3_{II}	3_{III}
Total Activity (U/mL)	0.740± 0.090	1.01± 0.065	0.888± 0.052	1.04± 0.003	0.712± 0.053	1.03± 0.032	1.36± 0.049	1.33± 0.058	1.086± 0.099

For all cultures except 3_I and 3_{III}, cultures expressed at 25˚C had higher total activity than those expressed at 20˚C. Induction at timepoints I (two doubling times) and II (two and a half doubling times) in all cases had the highest total activity, with the culture 3_{II} (at 25˚C) having the highest total activity (1.63 ± 0.058 U/mL, 47.9% of that reported in the initial paper).[19] From this point onward, all PaDa I-UPO expressions were carried out at 25 ˚C and induced at timepoint I.

The supernatants of the most active cultures (1_I, 1_{II}, 2_I, and 3_{II}) were combined, concentrated by filtration, and analyzed by SDS-PAGE (Figure 2.4).

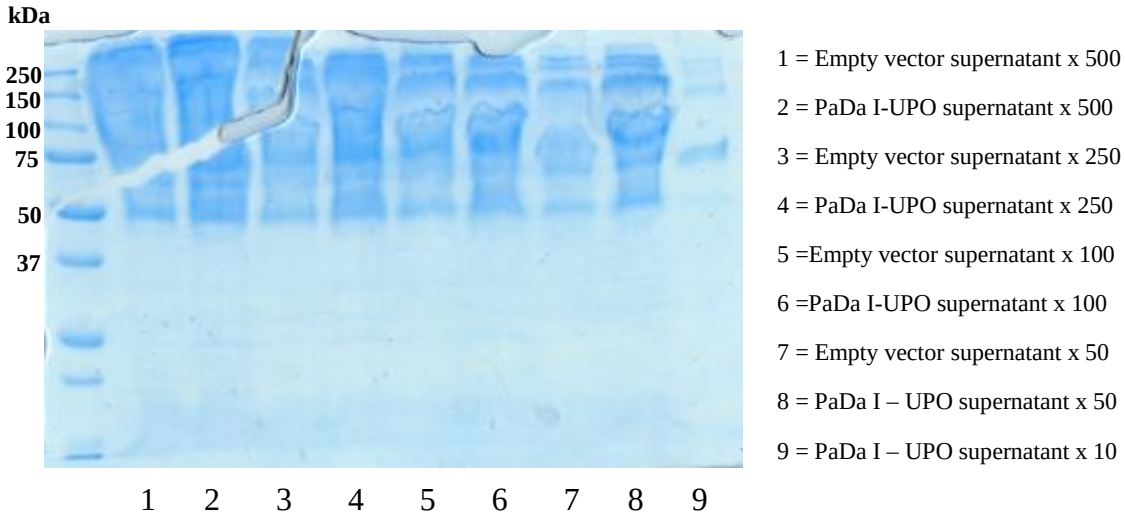


Figure 2.4 - SDS-PAGE analysis of *S. cerevisiae* culture supernatant at various concentrations. "x" in the legend above indicates level of concentration by filtration.

Despite higher levels of ABTS activity, there was no distinct protein band observed near PaDa I – UPO's expected molecular weight of 51.1 kDa (while a band is observed in most samples around 50 kDa, in Figure 7 the same band is observed in the empty vector control).

Attempts to purify PaDa I – UPO

In order to concentrate and partially purify the PaDa I-UPO protein, and thus observe a distinct PaDa I-UPO band, the PaDa I-UPO supernatant was fractionally precipitated with ammonium sulfate (Figure 2.5).

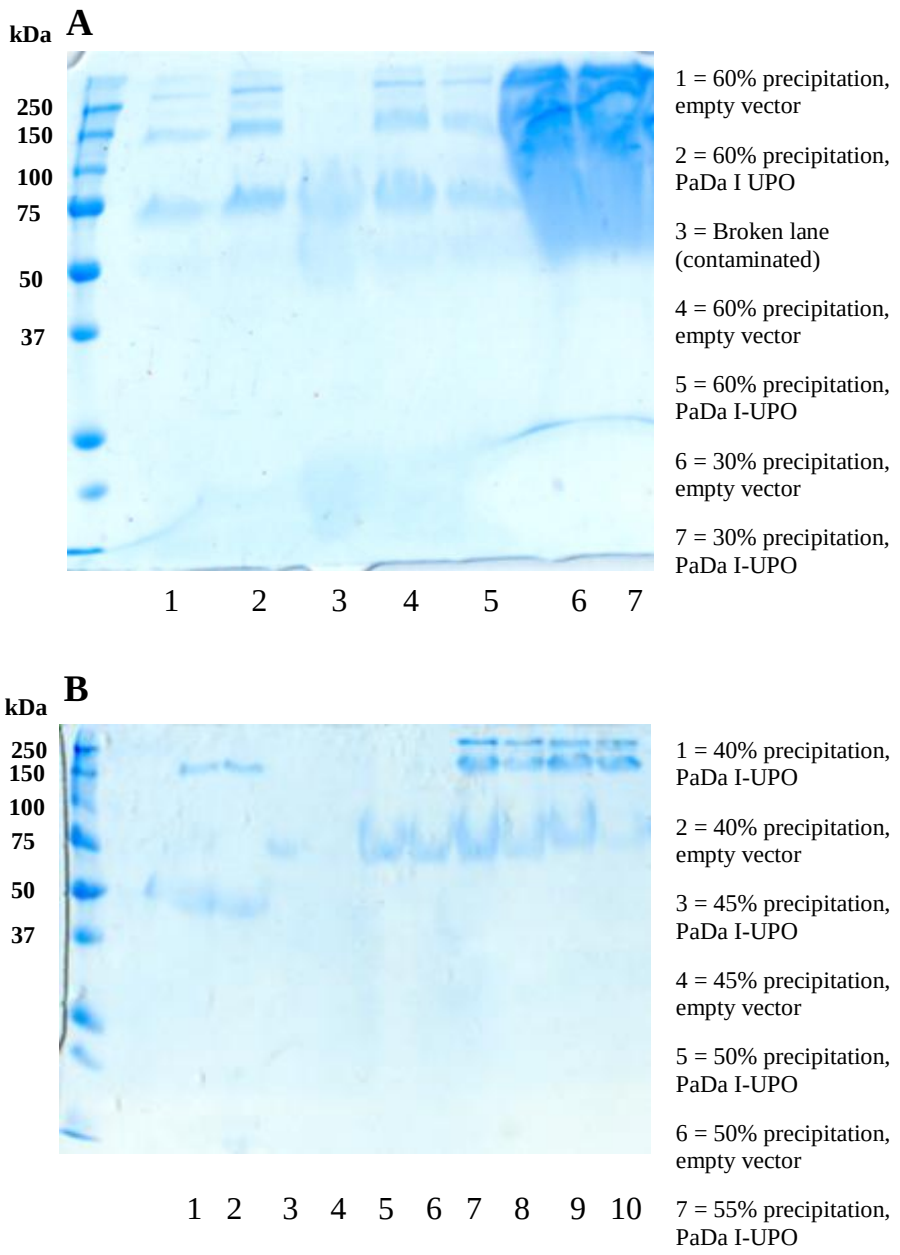


Figure 2.5 - SDS-PAGE analysis of PaDa I-UPO ammonium sulfate precipitation fractions. A) 30% precipitation followed by 60% precipitation and B) 30% precipitation (not shown) followed by successive precipitations to 60%. Percent of ammonium sulfate added is in w/v %. For a given experiment, all precipitations were performed successively on the same supernatant.

The resulting protein precipitates were too active to be measured with the ABTS assay, and were assayed against NBD instead (Table 2.5).

Table 2.5 - Activity of fractional ammonium sulfate precipitates on NBD. The number after the ± symbol is the standard deviation.

w/v percent of ammonium sulfate for precipitation	Unconcentrated supernatant	30%	40%	45%	50%	55%	60%
ΔOD425/min (NBD activity)	0.006 ±0.01	0.004 ±0.01	0.29 ±0.052	0.54 ±0.043	0.157 ±0.022	0.557 ±0.065	0.49 ±0.012

For comparison to total UPO activity as measured by ABTS, the unconcentrated supernatant used in these precipitations had a total activity of 0.942 ± 0.064 U/mL and the solution remaining after precipitation had a total activity of 0.00552 ± 0.00013 U/mL. From Figure 8A, it appears that much of the non-UPO protein in the supernatant is removed by 30% ammonium sulfate precipitation, while activity assays indicate that relatively little of the PaDa I-UPO protein is removed by 30% ammonium sulfate precipitation. Subsequent precipitations with higher levels of ammonium sulfate contained high concentrations of PaDa I-UPO (the 45% fraction is 90 times as concentrated as the supernatant, Table 2.5) and contained far less protein overall than non-precipitated supernatant (compare Figures 2.4 and 2.5). It also appears nearly all PaDa I-UPO was removed from the supernatant following 60% ammonium sulfate precipitation, as the remaining supernatant has low ABTS activity. However, the same protein bands appear to be present in both the empty vector control and PaDa I-UPO samples across all precipitation fractions, thus no distinct PaDa I-UPO band can be found

(Figure 2.5). It should be noted that Molina-Espeja et al. report an initial precipitation with 55% ammonium sulfate followed by a final cut of 85% ammonium sulfate, but experimentally the highest concentration of ammonium sulfate which could be dissolved in the supernatant was approximately 60% even with heating.

Following ammonium sulfate precipitation, ion exchange chromatography was used in an attempt to purify PaDa I-UPO. Despite using both dialysis and desalting columns (independently) to remove ammonium sulfate, the protein appeared not to bind to the strong cation exchange column. This was determined by ABTS assays, which established that all UPO activity was in the first several fractions to come off of the column, following loading and before elution with NaCl. The UPO activity containing fractions were compared by SDS-PAGE to empty vector control samples also subjected to the same FPLC method, resulting in Figure 2.6 below. Both empty vector control and PaDa I-UPO had the same visible bands, similar to those seen in Figure 2.5.

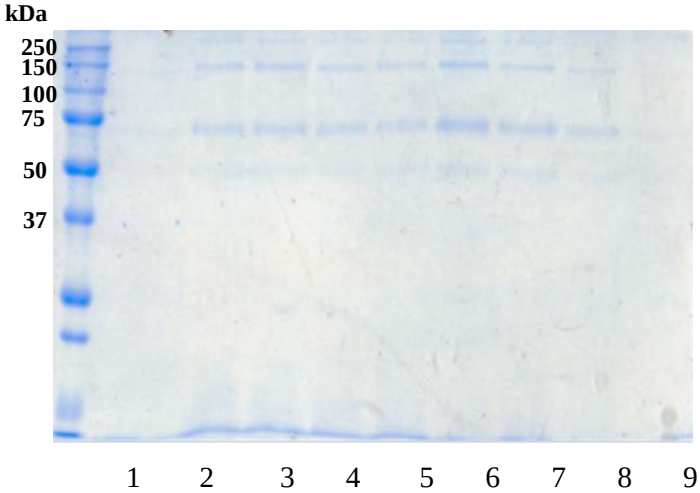


Figure 2.6 - SDS-PAGE analysis of PaDa I-UPO and empty vector control FPLC fractions. Lanes 1-5 are PaDa I –UPO fractions, while lanes 6-9 are empty vector control

Expression of PaDa I-UPO with vertebrate and *S. cerevisiae* Kozak sequences

In an attempt to increase PaDa I-UPO expression to a level observable by SDS-PAGE, the Kozak sequence on pESC-ura_PaDa_I was examined. Kozak sequences are conserved in the vicinity of the start codon in eukaryotic mRNA, and are recognized as the translational initiation site by the ribosome. The strength of the Kozak sequence is important for determining the amount of protein which will be synthesized by the ribosome from an mRNA transcript,[43] and thus the strength of the Kozak sequence in pESC-PaDa I could be adjusted to increase expression of PaDa I-UPO. The most common Kozak sequence identified in vertebrates is ACCATGG, [44] and this sequence is often used for heterologous expression in non-vertebrate hosts as it is very strong. Highly expressed genes in *S. cerevisiae* typically have the Kozak sequence AAAAAAATGTCT, making this a potentially useful sequence to test in heterologous expression along with the vertebrate Kozak sequence.[45]

In the area surrounding the start codon of PaDa I-UPO in pESC-ura_PaDa_I, no sequence similar to either the vertebrate or *S. cerevisiae* Kozak sequence was present. The surrounding sequence was CTAATGAAA (start codon underlined). In order to create a stronger Kozak sequence and increase production of PaDa I-UPO, site directed mutagenesis was used to produce six plasmids identical to pESC-ura_PaDa_I but with different Kozak sequences (Table 2.6).

Table 2.6 - pESC-ura PaDa I plasmids with altered Kozak sequences.

Plasmid	Sequence near start codon, altered Kozak sequence highlighted and start codon underlined
pESC-ura_PaDaI_A1	CGACTCAACCATGGAA
pESC-ura_PaDaI_A2	CGACTCAACCATGAAA
pESC-ura_PaDaI_A3	CGACAAAAAAATGTCT
pESC-ura_PaDaI_A4	CGACAAAAAAATGAAA
pESC-ura_PaDaI_A5	CGACTCAATAATGTCT
pESC-ura_PaDaI_A6	CGACTCAATAATGAAA

pESC-ura_PaDaI_A1 contains the full vertebrate Kozak sequence, which alters the first amino acid of PaDa I-UPO from lysine to glutamic acid. pESC-ura_PaDaI_A2 contains the vertebrate Kozak sequence but does not change lysine to glutamic acid. pESC-ura_PaDaI_A3 contains the full *S. cerevisiae* Kozak sequence, which changes the first amino acid of PaDa I-UPO from lysine to serine. pESC-ura_PaDaI_A4 contains the *S. cerevisiae* Kozak sequence, but does not change the first amino acid of PaDa I-UPO from lysine to serine. pESC-ura_PaDaI_A5 preserves the original surrounding sequence except for the third base in front of the start codon, which is switched from C to A, adenine at this position being the most crucial component of the Kozak sequence. pESC-ura_PaDaI_A5 also changes the first amino acid of PaDa I-UPO from lysine to serine in order to be closer to the *S. cerevisiae* Kozak sequence. pESC-ura_PaDaI_A6 is the same as pESC-ura_PaDaI_A5, but does not change the first amino acid of PaDa I-UPO from lysine to serine.

Plasmids pESC-ura_PaDaI_A4 and pESC-ura_PaDaI_A5 could not be transformed into BJ5465 *S. cerevisiae*, while plasmids pESC-ura_PaDaI_A1, pESC-ura_PaDaI_A2, pESC-ura_PaDaI_A3, and pESC-ura_PaDaI_A6 were transformed into BJ5465 *S. cerevisiae* and expressed. Supernatants were assayed for ABTS activity (Table 2.7), and then concentrated by filtration for SDS-PAGE analysis (Figure 2.7).

Table 2.7 - Total activities (determined by ABTS assay) of PaDa I-UPO supernatants expressed with modified Kozak sequences. The number after the ± symbol is the standard deviation.

Plasmid contained in culture	Total activity of supernatant (U/mL)
pESC-ura_PaDaI_A1	No activity
pESC-ura_PaDaI_A2	1.47 ± 0.0482
pESC-ura_PaDaI_A3	No activity
pESC-ura_PaDaI_A6	1.58 ± 0.040

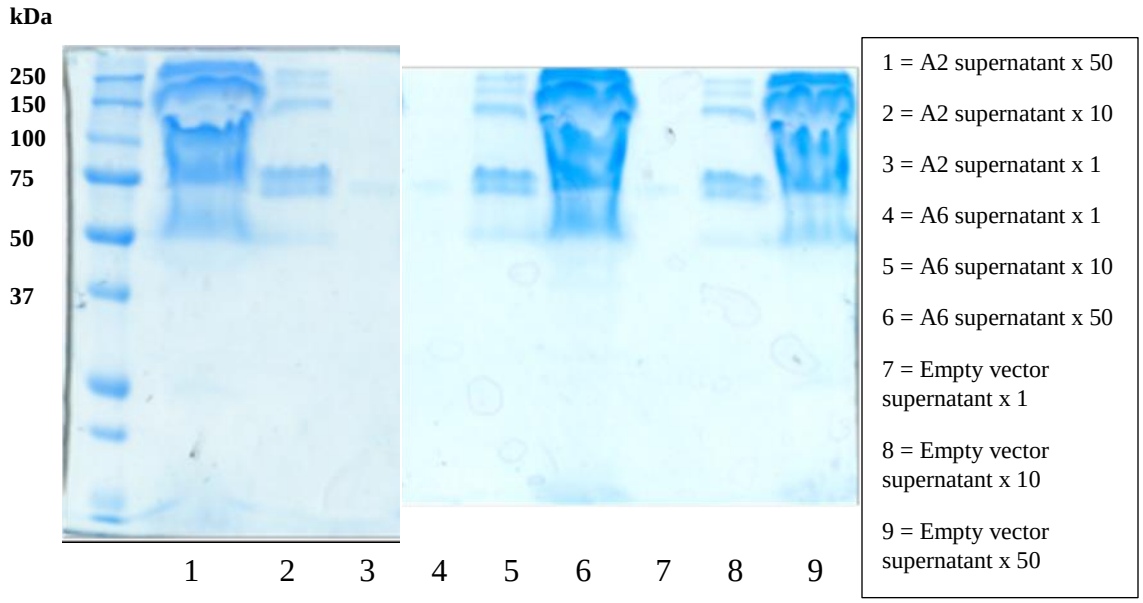


Figure 2.7 - SDS-PAGE analysis of concentrated expression supernatant from cultures containing pESC-ura PaDa I A2 and A6. The "x" in the legend above indicates level of concentration by filtration. Note: the figure above is composed of two independent gels which were run together.

Only two of the expression cultures, those expressing PaDa I-UPO with Kozak sequences A2 and A6, were active, and these did not have significantly higher activity than PaDa I-UPO expressed with the native sequence (see Table 2.7). Both of the plasmids which altered the first amino acid of PaDa I-UPO from lysine to glutamic acid or serine had no activity, indicating that the presence of lysine in this position is crucial to PaDa I-UPO's activity. As the supernatants from the two active cultures did not contain a distinct (i.e. not also seen in the empty vector control) protein band corresponding to PaDa I's expected molecular weight of 51.1 kDa, it appears the use of modified Kozak sequences did not increase PaDa I-UPO expression.

Modification of terpenes by PaDa I-UPO

After repeated purification attempts, concentrated (but unpurified) supernatant with peroxidative and peroxygenase activity was assayed by GC/MS for activity against various terpenes. As *A. aegerita* UPO has previously been shown to modify limonene,[23] this substance was assayed first (see Figure 2.8).

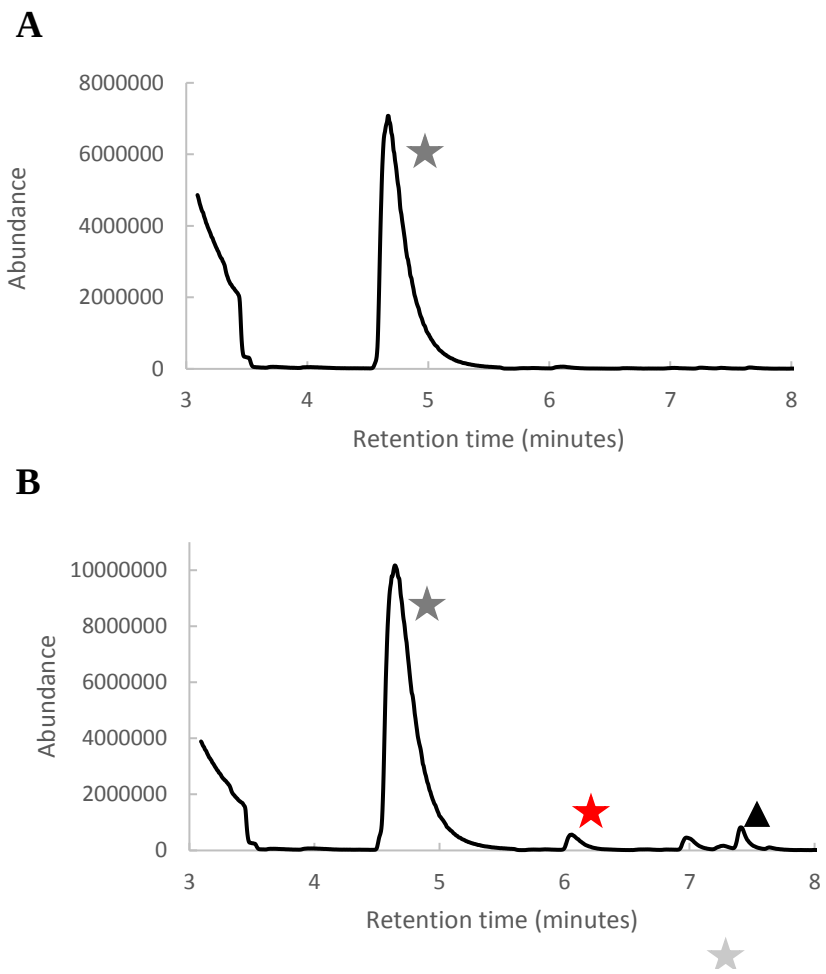


Figure 2.8 - GC/MS analysis of PaDa I – UPO limonene reactions. A) reaction of limonene with empty vector control supernatant and **B)** reaction of PaDa I-UPO supernatant with limonene. A grey star indicates limonene, a red star indicates limonene epoxide, while a black triangle indicates carveol. All peak assignments were determined by comparison with the National Institute of Standards and Technology (NIST) molecular database.[46]

As reported in the literature for *A. aegerita* UPO, the PaDa I-UPO containing supernatant converted limonene to limonene epoxide and carveol.[23] To determine whether PaDa I-UPO can modify sesquiterpene compounds, the PaDa I-UPO supernatant was assayed against Δ6-protoilludene, which was provided by co-reaction with Δ6-protoilludene synthase. A peak with an m/z value of 220 was observed in both the liquid and headspace of the Δ6-protoilludene assay. As Δ6-protoilludene has an m/z value of 204, this peak could be a Δ6-protoilludene derivative with an additional oxygen atom (MW of 16 g/mol). No peak with m/z of 220 was observed in the absence of Δ6-protoilludene synthase, indicating Δ6-protoilludene must be present for the putatively modified compound to appear. (Figure 2.9 and 2.10).

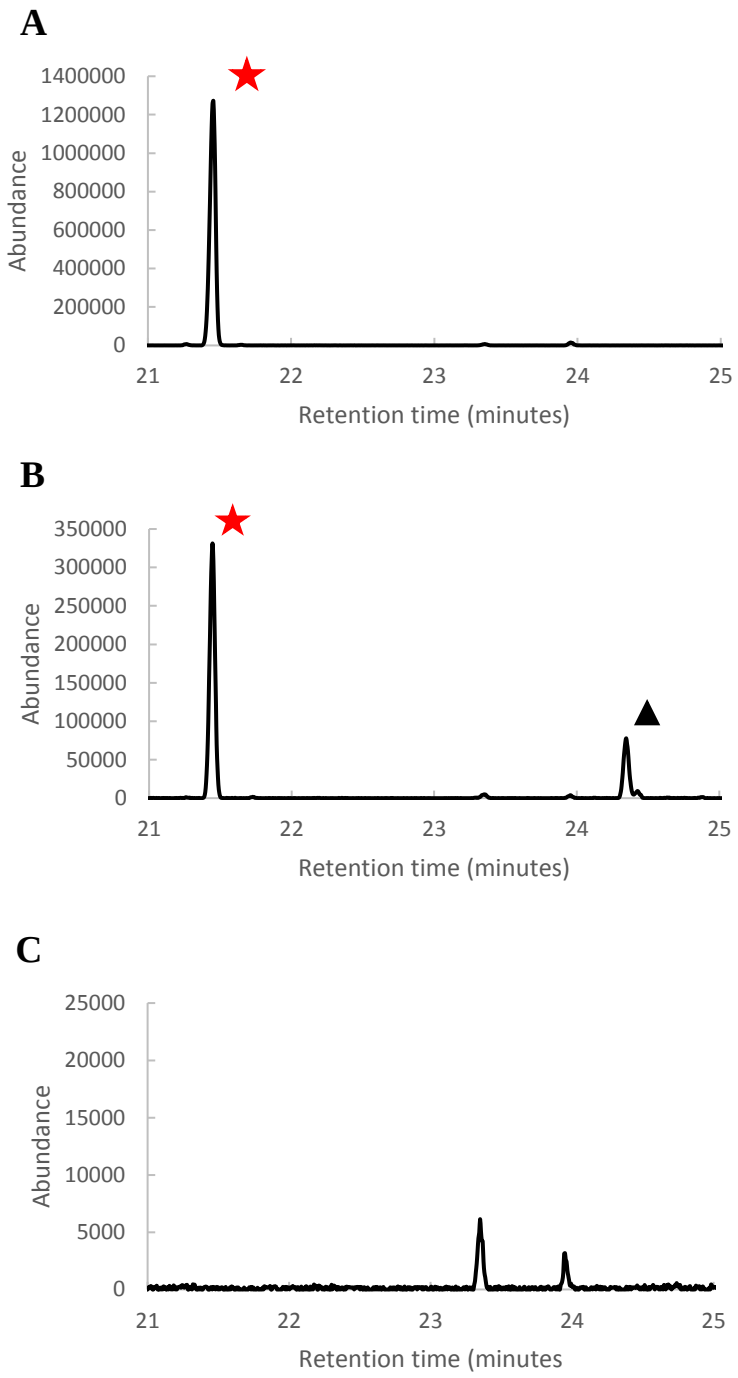


Figure 2.9 - GC/MS analysis of Δ6-protoilludene and PaDa I-UPO reaction, liquid fraction. A) Δ6-Protoilludene synthase and empty vector supernatant B) Δ6-Protoilludene synthase and PaDa I-UPO supernatant C) FPP and PaDa I-UPO. Red stars indicate Δ6-protoilludene (as determined by NIST database),[46] and black triangles represent the putatively modified product with a molecular weight of 220 g/mol.

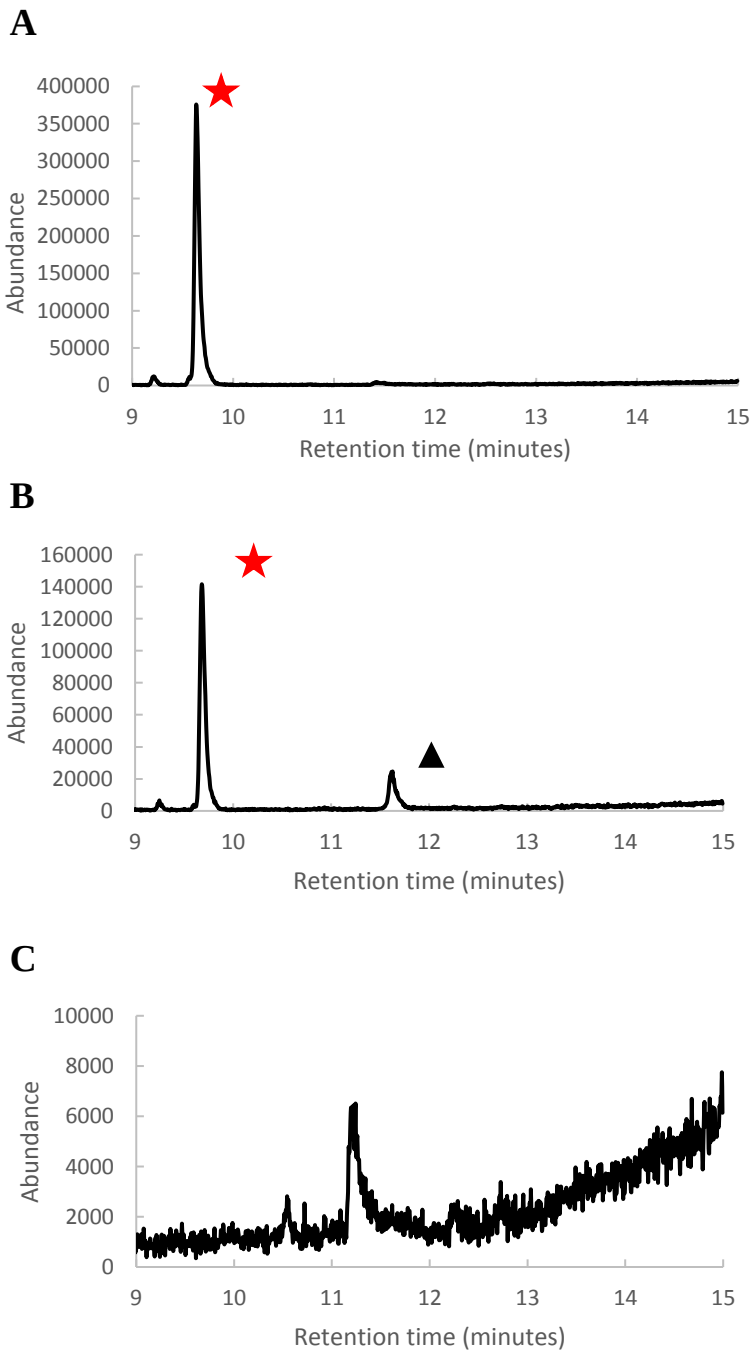


Figure 2.10 - GC/MS analysis of Δ6-protoilludene and PaDa I-UPO reaction, volatile headspace. A) Δ6-Protoilludene synthase and empty vector supernatant B) Δ6-Protoilludene synthase and PaDa I-UPO supernatant C) FPP and PaDa I-UPO. Red stars indicate Δ6-protoilludene (as determined by NIST database),[46] and black triangles represent the putatively modified product with a molecular weight of 220 g/mol.

To determine if the reaction is dependent on H_2O_2, PaDa I – UPO was also assayed against Δ6-protoilludene at high (5x) H_2O_2 concentration and without H_2O_2 (see figure 2.11).

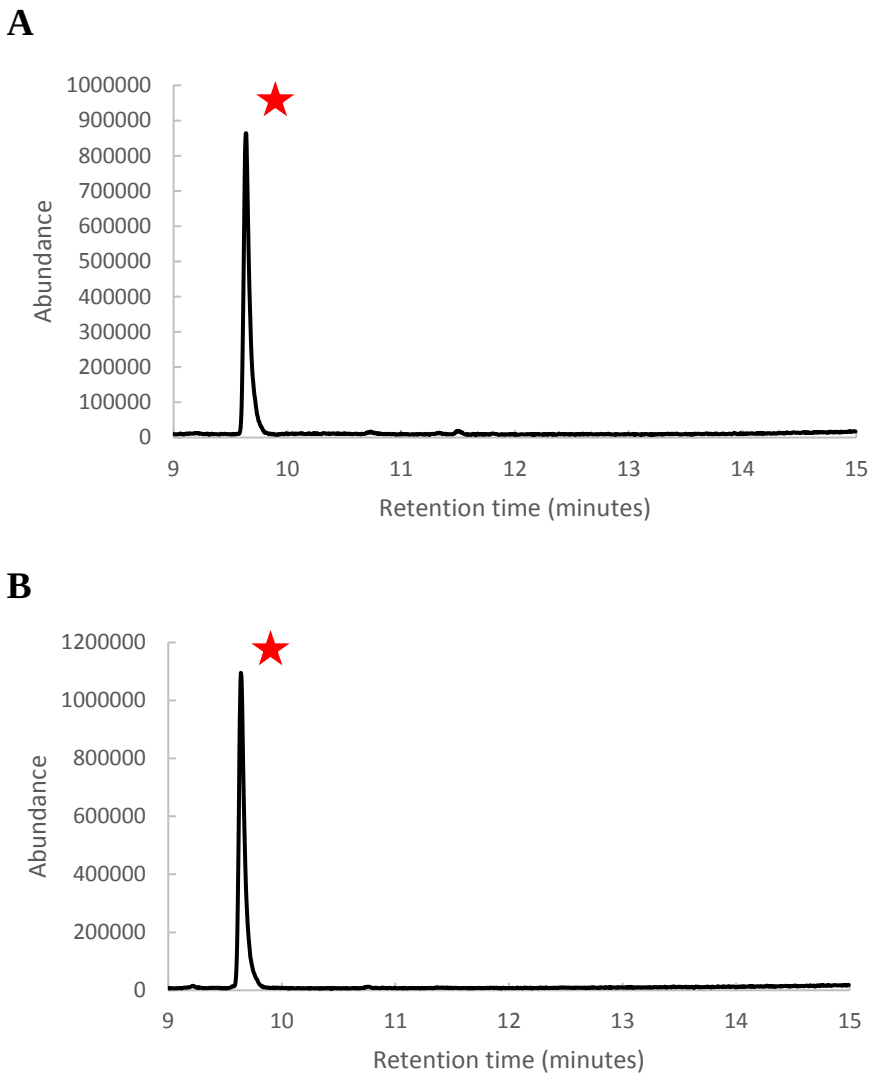


Figure 2.11. GC/MS analysis of Δ6-protoilludene and PaDa I-UPO reaction, volatile headspace. A) Reaction with 5x increased H_2O_2 B) Reaction without H_2O_2. Red stars indicate Δ6-protoilludene (as determined by NIST database)

As expected for hydrogen peroxide dependent UPO, no modified product is observed in the absence of H_2O_2. Increasing the level of H_2O_2 in the reaction also resulted in an absence of modified product. This is in accordance with established literature, as increased H_2O_2 has been shown to inactivate *A. aegerita* UPO, most likely by the production of hydroxyl radicals which react with the heme of the active site and produce biliverdin.[47] Even for reactions with the level of H_2O_2 used for the GC/MS assays shown above, this heme inactivation or "heme bleaching," along with the intrinsic catalase activity of UPO, could account for the fact that the modified Δ6-protoilludene peak is considerably smaller than the unmodified peak.

In addition, PaDa I – UPO was assayed against another sesquiterpene, valencene, using the same procedure used for limonene. Initially three sesquiterpenes, caryophyllene, humulene, and valencene were to be assayed, but the caryophyllene and humulene in our possession had already become oxidized in the bottle. No activity was observed for PaDa I – UPO against valencene (see supplementary materials).

Conclusion

Based on ABTS and NBD assays for peroxidative and peroxygenase activity, respectively, it appears that PaDa I-UPO was successfully expressed in *S. cerevisiae*. The highest activity against ABTS achieved in any culture was 1.63 ± 0.058 U/mL, 47.9% of that reported in the literature.[19] In addition, culture supernatant modified limonene to limonene epoxide and carveol (as shown in the literature),[23] and appears to have activity against Δ6-protoilludene. According to GC/MS analysis, a compound with a molecular

weight of 220 g/mol is produced, which is consistent with the addition of an oxygen atom to Δ6-protoilludene. Despite this activity, SDS-PAGE analysis revealed that the PaDa I-UPO supernatant and empty vector control supernatant had the same visible protein bands. Altering expression conditions such as temperature and time of induction, and adding Kozak sequences based on the vertebrate and *S. cerevisiae* sequences did not significantly increase PaDa I-UPO production, or cause a distinct PaDa I-UPO protein band to appear. No band was revealed either by concentration and purification with ammonium sulfate precipitation or ion exchange chromatography.

Chapter 3

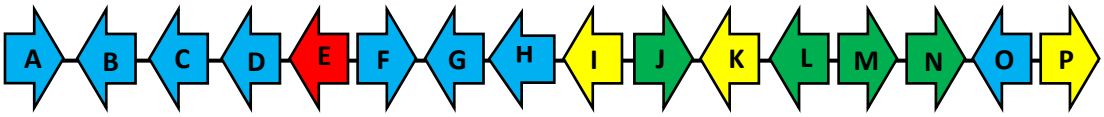
Expression of fungal flavin binding enzymes in *E. coli*

Introduction

Flavin dependent monooxygenases (FMOs) catalyze the transfer of an atom of molecular oxygen to a substrate molecule, while the other oxygen atom is reduced to water. In nature, FMOs are involved in catabolism, hormone biosynthesis, vitamin and antibiotic production, and defense .[48] They are known to catalyze a variety of reactions, including hydroxylation, epoxidation, Baeyer–Villiger oxidation, sulfoxidations, and halogenations. These oxidation reactions would be either impossible or very difficult to achieve by organic chemical synthesis.[49] Because of this, and because of their high enantio- and regio- selectivity, FMOs have attracted the attention of the pharmaceutical, fine-chemical and food industries.

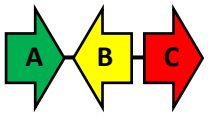
The gene cluster around Omp7 protoilludene synthase in *Omphalatus olearius* contains an enzyme called Omp7a,[16] which has been identified through bioinformatics to be a FAD binding oxidoreductase. In addition, Omp7a has been co-expressed with the Δ6-protoilludene synthase Omp7 in *S. cerevisiae*, and appears to have some activity, producing a non-volatile compound that degrades in the GC/MS (unpublished data). The Stehi 7 protoilludene synthase gene cluster in *Stereum hirsutum* contains a number of enzymes[17] which have been identified by bioinformatics as potential scaffold modifiers. These include the FAD binding oxidoreductases FAD1 and FAD2, the GMC (glucose-methanol-choline oxidase) oxidoreductase GMC2, and the reductase RED1.

Note that the GMC superfamily are flavoprotein oxidoreductases.[50] In fungi, genes involved in secondary metabolite biosynthesis are clustered together on the genome.[51] Thus the presence of flavin binding genes in Δ6-protoilludene synthase gene clusters indicate that the flavin binders are likely part of terpenoid biosynthesis. Presumably, they modify the Δ6-protoilludene scaffold. As the CYPs of these gene clusters have proven difficult to isolate, it is worthwhile attempting to isolate other Δ6-protoilludene modifying enzymes for use in an *in vitro* terpenoid biocatalytic pathway.



A = Glycosyl hydrolase	I = CYP (cytochrome P450)
B = Acyl-CoA transferase	J = Reductase (RED1)
C = Outer membrane protein	K = CYP
D = Aspartate aminotransferase	L = GMC oxidoreductase (GMC2)
E = Stehi7 protoilludene synthase	M = FAD binding oxidoreductase (FAD1)
F = Aldo-keto reductase	N = FAD binding oxidoreductase (FAD2)
G = GMC oxidoreductase	O = Aldo-keto reductase
H = Major facilitator superfamily (MFS) transporter	P = CYP

Figure 3.1. Stehi7 biosynthetic gene cluster. [17] Stehi7 protoilludene synthase is pictured above in red, CYPs in the gene cluster are in yellow, flavin binding enzymes investigated in this study are in green, all other genes of the cluster are in blue.



A = FAD-binding oxidoreductase (Omp7a) B = CYP C = Omp7 protoilludene synthase

Figure 3.2. Omp7 biosynthetic gene cluster.[16] Omp7 protoilludene synthase is in red, Omp7a (investigated in this study) is in green, and CYPs are yellow.

The key to the oxygenase activity of FMOs is the flavin prosthetic group itself, which during the reaction becomes a flavin C4a-oxygen adduct and transfers oxygen to the substrate.[52] In most FMOs the flavin cofactor is non-covalently, but tightly, bound to the apoenzyme and the two can be disassociated.[49] However, some flavoenzymes are attached to their flavin cofactor by one or two covalent linkages.[53]

The general mechanism of the flavin dependent monooxygenase reaction cycle is shown below. The reaction of the reduced flavin with oxygen results in formation of the nucleophilic flavin C4a-peroxide species or the electrophilic C4a-hydroperoxide species, which then react with either an electrophilic or nucleophilic substrate molecule, respectively. Once oxygen is transferred to the substrate hydroxyflavin forms. Loss of water then produces oxidized flavin, which is reduced in a critical step to restart the cycle.[48]

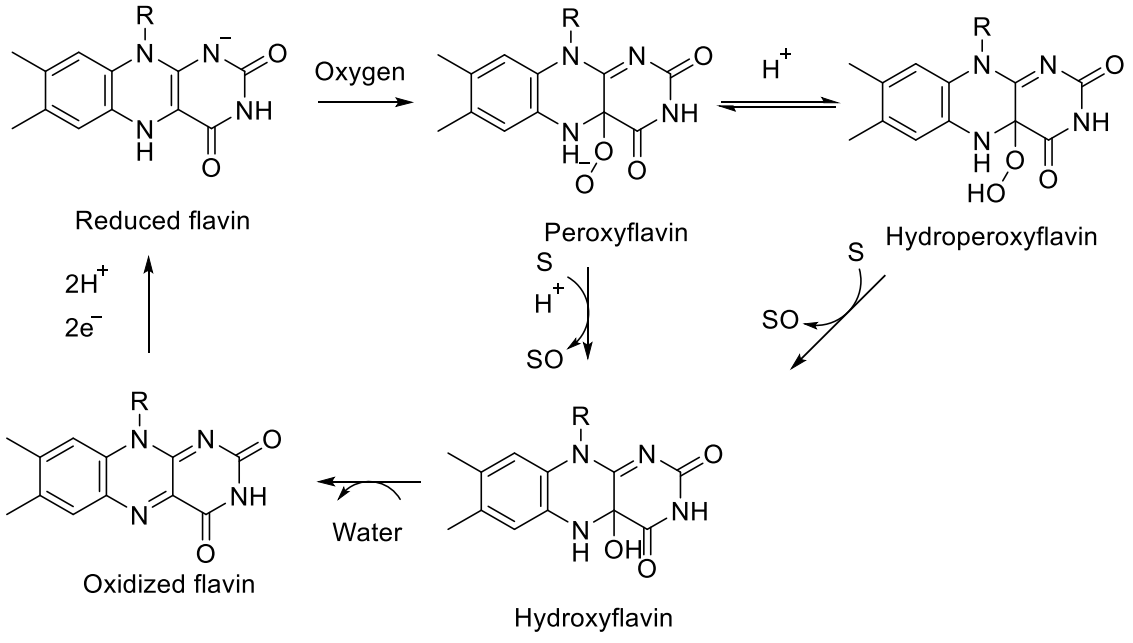


Figure 3.3. The general mechanism of the flavin dependent monooxygenase reaction cycle[48]

Flavoenzymes can be divided into groups based on whether they utilize NAD(P)H as an external electron donor, use a flavin reductase partner to reduce flavin, or reduce the flavin cofactor through substrate oxidation.[48] As the largest group of FMOs utilize a flavin reductase to reduce flavin, the *E. coli* flavin reductase (Fre) will overexpressed in *E. coli* so that it can be used *in vitro*.

Materials and Methods

Transfer of flavin binding enzymes and *E. coli* Fre to pCuminBB plasmids.

In order to obtain the *E. coli* flavin reductase (Fre) gene, a Top10 *E. coli* culture was grown overnight at 37 ˚C and 220 RPM. The *E. coli* genomic DNA was extracted with a Promega Wizard® DNA purification kit. Primers to amplify *E. coli* Fre from gDNA were based on the *E. coli* Fre sequence reported by Spyrou et al.,[54] while the other flavin binding enzymes were amplified from plasmids in the Schmidt-Dannert collection (see Table 3.1). FAD1, FAD2, GMC2, RED1, Omp7a, and *E. coli* Fre were cloned into pCuminBB plasmids, Fre with an N-terminal histidine tag and the other enzymes with C-terminal histidine tags. Cloning was performed by Hifi assembly (using the online NEBuilder® assembly tool to design primers, the NEBuilder® HiFi DNA Assembly kit and NEB's protocols), or with restriction digestion and ligation (see Table 3.2). Primers were ordered from Integrated DNA Techologies. Thermo Fisher Scientific Top10® cells were used for plasmid production during cloning, GoGreen® Taq polymerase from Promega was used to screen colonies for desired genes, and sequencing was performed at the University of Minnesota Genomics Center. pCuminBB plasmids containing FAD-binding enzymes were transformed into C2566 *E. coli* cells for protein expression.

Table 3.1 – Plasmids and strains used in Chapter 3

Plasmid name	Gene and promoter information	Selectable marker	Source
pCuminBB-ctH6	Contains cumate promoter, histidine tag for protein C-terminus	Ampicillin resistance	From Maureen Quin (CSD lab)
pUCBB-FAD1	Contains FAD1 gene	Ampicillin resistance	From Christopher Flynn (CSD lab)
pUCBB-FAD2	Contains FAD2 gene	Ampicillin resistance	From Christopher Flynn (CSD lab)
pCR-Blunt-GMC2	TOPO™ (Thermofisher Scientific) vector containing GMC2 gene	Kanamycin and zeocin resistance	From Christopher Flynn (CSD lab)
pCR-Blunt-RED1	TOPO™ vector containing RED1 gene	Kanamycin and zeocin resistance	From Christopher Flynn (CSD lab)
pESC-His-Omp7a	Contains Omp7a	Ampicillin resistance, HIS3* Marker	From Sarah Perdue (CSD lab)
pCuminBB-cHis-FAD1	Contains ct-His**-FAD1 gene under control of cumate promoter	Ampicillin resistance	This study
pCuminBB-cHis-FAD2	Contains ct-His-FAD2 gene under control of cumate promoter	Ampicillin resistance	This study
pCuminBB-cHis-GMC2	Contains ct-His-GMC2 gene under control of cumate promoter	Ampicillin resistance	This study
pCuminBB-cHis RED1	Contains ct-His-RED1 gene under control of cumate promoter	Ampicillin resistance	This study
pCuminBB-cHis Omp7a	Contains ct-His-Omp7a gene under control of cumate promoter	Ampicillin resistance	This study
pCuminBB-ntH6	Contains cumate promoter, histidine tag for protein N-terminus	Ampicillin resistance	From Maureen Quin (CSD) lab
pCuminBB-nHis-E.coli-Fre	Contains nt-His***-Fre (*E.coli* flavin reductase) gene under control of cumate promoter	Ampicillin resistance	This study
pCuminBB-nHis-E.coli-Fre ATG-onlyb4his	Contains nt-His-Fre gene under control of cumate promoter. Native start codon of Fre removed	Ampicillin resistance	This study
Strain	Description	Source	
Top10	Chemically competent *E. coli* strain used for plasmid production	ThermoFisher Scientific	
C2566	*E. coli* strain used for protein expression	New England Biolabs (NEB)	

*HIS3 allows growth on minimal media without added histidine.

**ct-His indicates the protein has a C-terminal histidine tag

***nt-His indicates the protein has an N-terminal histidine tag

Table 3.2 – Primers used in Chapter 3

Primer name	5' – 3' Sequence	Function
FAD1 F*	GGATCCATATGCCTGGCAAACTCAATG	Amplifies FAD1 from pUCBB-FAD1, amplicon is cut with NdeI and NotI restriction enzymes, ligated into cut pCuminBB-ctH6 vector to produce pCuminBB-cHis-FAD1
FAD1 R	GTAGTGCGGCCGCGGCAGCAACAGGCCTGTC	
FAD2 F	GGATCCATATGGCATCGGAAAAGAACTCTC	Amplifies FAD2 from pUCBB-FAD2, amplicon is cut with NdeI and NotI restriction enzymes, ligated into cut pCuminBB-ctH6 vector to produce pCuminBB-cHis-FAD2
FAD2 R	CCAGTGCGGCCGCGAAATGCCATCCTCCTGC	
GMC2 F	CTGATGTCGACATGTCTCTTTCACCGTCCTTG	Amplifies GMC2 from PCR-Blunt-GMC2, amplicon is cut with SalI and NotI restriction enzymes, ligated into cut pCuminBB-ctH6 vector to produce pCuminBB-cHis-GMC2
GMC2 R	CCAGTGCGGCCGCTAGCTCTCCTGCTTGACGTG	
A11F	AGCAGCGGCCATCATCAT	Linearizes pCuminBB-ctH6 for HiFi assembly of RED1 and Omp7a
A7R	GGATCCAGATCCCTCCTTC	
A9F	CGAAGGAGGGATCTGGATCCATGTCCTCTCACTCTGCAG	Amplifies RED1 from PCR-Blunt-RED1, for HiFi assembly to produce pCuminBB-cHis RED1
A13R	ATGATGGCCGCTGCTGCTGCCGCGCGGCACCAGGGATCCTGACAAAGCGATGTCGAC	
A10F	CGAAGGAGGGATCTGGATCCATGTCTGCTCCTGCTTCTTTCAC	Amplifies Omp7a from pESC-His-Omp7a for HiFi assembly to produce pCuminBB-cHis Omp7a
A14R	ATGATGGCCGCTGCTGCTGCCGCGCGGCACCAGGGATCCCACGGAAATCGGGCACCA	
A21F	ATATGGGCGGTTGATAAGATATC	Linearizes pCuminBB-ntH6 for HiFi assembly of *E. coli* Fre
A21R	GGATCCGCTGCCGCGCGG	
A22F	CCTGGTGCCGCGCGGCAGCGGATCCATGACAACCTTAAGCTGTAAAG	Amplifies *E. coli* Fre from *E. coli* genomic DNA for HiFi assembly to produce pCuminBB-nHis-E.coli-Fre
A22R	TCGATATCTTATCAACCGCCCATATTCAGATAAATGCAAACGC	
A23F	CCTGGTGCCGCGCGGCAGCGGATCCACAACCTTAAGCTGTAAAGTG	Amplifies E. coli Fre from E. coli genomic DNA for HiFi assembly to produce pCuminBB-nHis-E.coli-Fre- ATG-onlyb4his
A23R	TCGATATCTTATCAACCGCCCATATTCAGATAAATGCAAACGC	

* = F indicates forward primer, R indicates reverse primer

Expression of flavin-binding enzymes in *E. coli*

A C2566 *E. coli* colony containing the desired plasmid (ampicillin resistant) was picked from a selective plate and used to inoculate a 4 mL culture of lysogeny broth (LB) containing 100 μg/mL of ampicillin. LB contained 10 g of Bacto-tryptone, 5 g of yeast extract, and 5 g of NaCl per liter. This culture was grown overnight at 37 ˚C and 220 RPM. The following morning an amount of starter culture equal to $1/100^{th}$ of the expression culture's volume was used to inoculate an LB expression culture containing 100 μg/mL of ampicillin. Unless otherwise noted, expression cultures were 4 mL. Expression cultures were grown until they reached an OD600 of 0.4-0.5, upon reaching which protein production was induced by addition of cumate (ρ-isopropylbenzoate) from a 50 mM cumate stock in ethanol.

The expression cultures were grown at 37 ˚C (unless otherwise noted) and grown overnight (approximately 16 hours), unless otherwise noted. To harvest the cells and test protein expression level, 1 mL of each expression culture was taken, spun down at 13000 rpm for 1 minute, the supernatant was removed, and the resulting cell pellet was resuspended in 300 μL of Bugbuster® protein extraction reagent for lysis. The lysis reactions were left at room temperature (with occasional shaking) for 20 minutes, and then 15 μL of the lysis solution was removed and labeled the "total protein fraction". The remaining solution was spun down at 13000 rpm for 5 minutes, resulting in clear supernatant and a cell pellet. 15 μL of the clear supernatant was removed and labeled the

"soluble protein fraction." The protein expression level was determined by SDS-PAGE analysis, and with staining in Coomassie Biosafe® stain overnight.

Co-expression of flavin binding enzymes with chaperone proteins in *E. coli*.

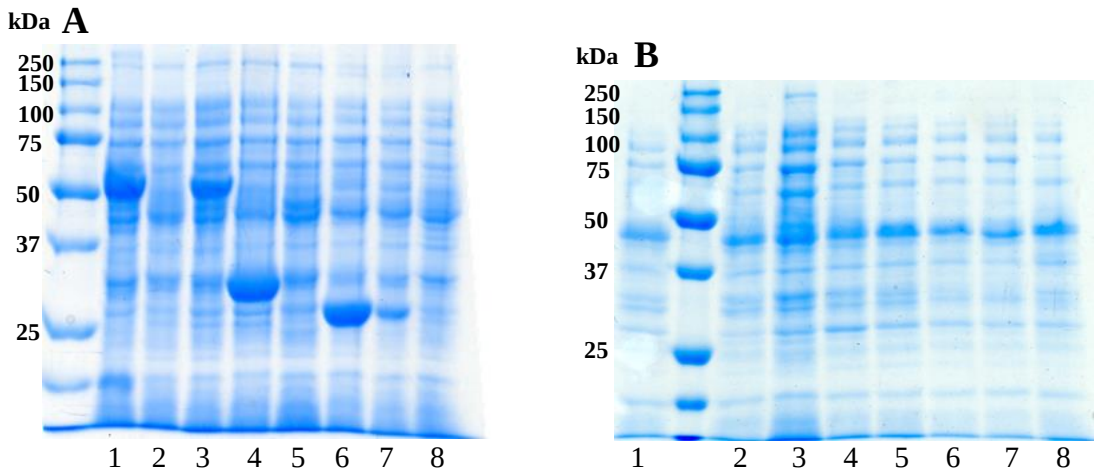
FAD1, GMC2, and RED1 were each co-expressed in C2566 *E. coli* cultures with the molecular chaperones of the Takara Chaperone Set. Takara has included a very detailed protocol in their Chaperone Plasmid Set Product Manual (Cat #3340) for how to perform coexpression of the chaperones and a target protein. This protocol was adapted to the 4 mL expression performed with the flavin binding enzymes (described in the section immediately above) and performed at 37 ˚C.

Results and Discussion

Expression of flavin-binding enzymes with varying inductant concentrations

Genes encoding the flavin-binding enzymes FAD1, FAD2, GMC2, RED1, and Omp7a were obtained from plasmids in the Schmidt-Dannert plasmid collection (see Table 3.2), while the gene encoding *E. coli* FAD-reductase (Fre) was amplified from extracted *E. coli* genomic DNA. The *E.coli* Fre gene was inserted into a pCumin plasmid with an N-terminal histidine tag, while the other flavin binding enzymes were inserted into pCumin plasmids with C-terminal histidine tags. pCumin plasmids use the cumate inducible system, which contains a strong constitutive promoter and an operator (CuO) that is bound by the repressor CymR. Upon addition of the inductant cumate (ρ-isopropylbenzoate), repression is lifted.[55] As the N-terminal histidine tag contains its own start codon, it is possible that the ribosome could start translation at the start codon of *E. coli* Fre proper and not include the histidine tag. Because of this, plasmids containing N

terminally histidine tagged Fre with and without a separate start codon before the Fre gene were made. Expressions were performed at 37˚C with 50 µM cumate, and both soluble and total (containing soluble and insoluble) protein fractions were analyzed by SDS-PAGE (Figure 3.4 and Table 3.3).



1 = FAD1, 2 = FAD2, 3 = GMC2, 4 = RED1, 5 = Omp7a, 6 = *E. coli* Fre w/ATG, 7 = *E. coli* Fre w/out ATG, 8 = Empty pCuminBB-cHis (negative control)

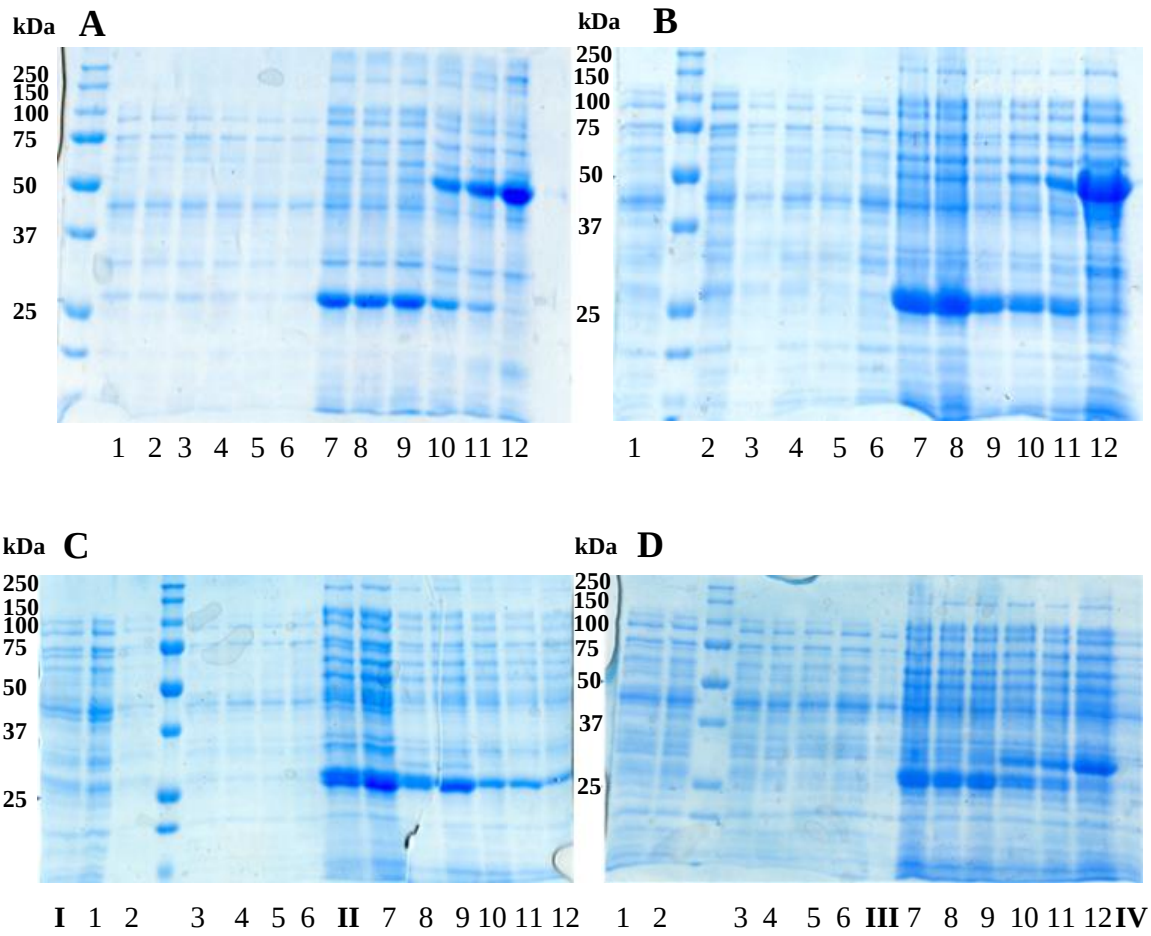
Figure 3.4. SDS-PAGE analysis of flavin-binding enzyme expression at 37 ˚C and 50 µM cumate A) total and B) soluble protein fractions. Refer to Table 3.3 for expected molecular weights of FAD binding enzymes. Note, these expressions were performed in 500 mL cultures.

Table 3.3. Expected molecular weights of FAD binding enzymes. Calculated using ExPASy® program

FAD binding enzyme	GMC2	FAD1	FAD2	RED1	Omp7a	*E. coli* Fre
Expected Molecular Weight	68.5 kDa	52.2 kDa	57.3 kDa	32.6 kDa	49.8 kDa	26.2 kDa

Protein bands corresponding to FAD1, GMC2, RED1, and both *E. coli* Fre with and without its own start codon are visible at approximately the expected molecular weights in the insoluble fraction, but not in the soluble fraction, indicating that these proteins are expressed but insoluble. The insoluble protein band for *E. coli* Fre with its own start codon is considerably stronger than that seen when only the start codon before the N terminal histidine tag is present, indicating that the presence of a start codon immediately before Fre proper increases expression. A faint band is seen for Omp7a between 37 and 50 kDa in the insoluble fraction, indicating this protein may possibly be expressed insolubly, but very poorly. Both soluble and insoluble fractions for FAD2 expression contained no bands distinct from the empty vector control, indicating this protein is not expressed.

The expression of proteins as insoluble inclusion bodies is often due to protein synthesis being faster than protein folding, resulting in high concentrations of unfolded protein which aggregate due to the exposure of normally buried hydrophobic residues.[41] Slowing down protein synthesis can result in increased solubility, and this is often accomplished by decreasing expression temperature and decreasing the concentration of inductant[41,42] As cumate is the inductant for the flavin-binding enzyme expression system, the strongly expressed proteins (FAD1, GMC2, RED1, and *E. coli* Fre) were expressed with varying levels of cumate to determine if a certain concentration would yield soluble protein (Figure 3.5).



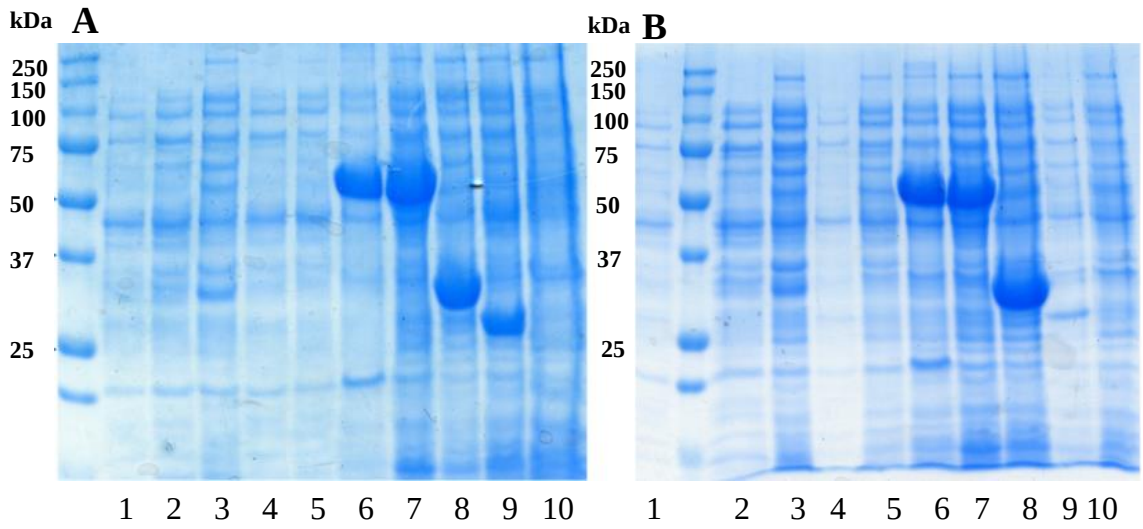
1 = 1.25 µM (soluble), 2 = 5 µM (soluble), 3 = 10 µM (soluble), 4 = 20 µM (soluble), 5 = 30 µM (soluble), 6 = 50 µM (soluble), 7 = 1.25 µM (total), 8 = 5 µM (total), 9 = 10 µM (total), 10 = 20 µM (total), 11 = 30 µM (total), 12 = 50 µM (total)
I = 0 µM (soluble), **II** = 0 µM (total), **III** = Empty pCuminBB-cHis 50 µM (soluble), **IV** = Empty pCuminBB-cHis 50 µM (total),

Figure 3.5. SDS-PAGE analysis of flavin-binding enzyme expression at 37 ˚C and a range of cumate concentrations. A) FAD1 expression B) GMC2 expression C) *E. coli* Fre expression D) RED1 expression. See Table 3.3 for expected molecular weights of flavin-binding enzymes. Concentrations of cumate used in each expression are shown in the legend above, and whether the protein fraction is soluble or total is indicated in parenthesis.

For FAD1, GMC2, and RED1, significant protein expression in the total protein fraction occurs at 20 μM cumate, increases with increasing cumate concentration, and is strongest at 50 μM cumate. However, FAD1, GMC2, and RED1 do not express in soluble form at any cumate concentration. In contrast, *E. coli* Fre appears to express at 0 μM cumate, with expression apparent in all cultures (even those not containing the Fre pCumin expression plasmid), and expression decreases as cumate concentration is increased.

Variation of expression temperature and induction time

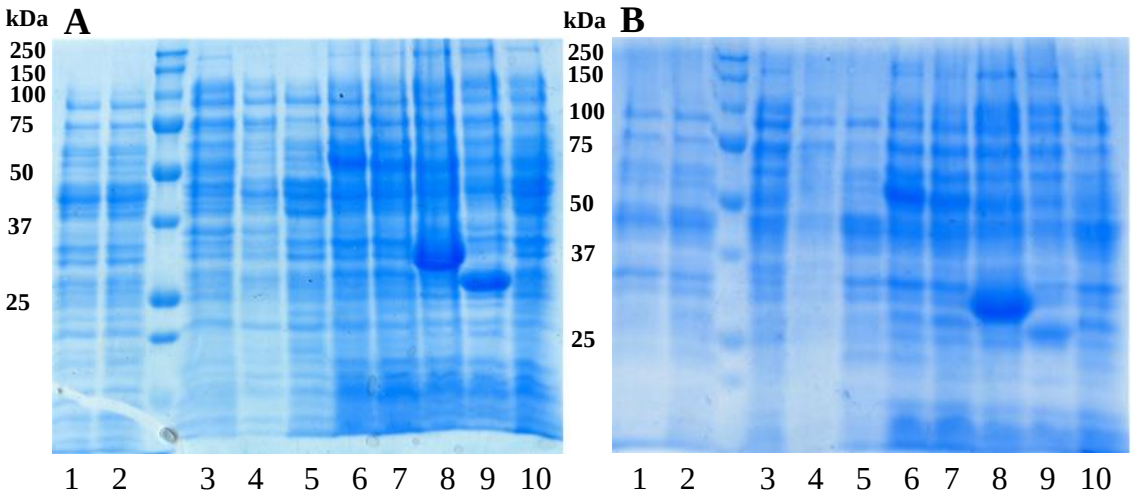
As expression of flavin-binding enzymes is strongest at 50 μM cumate, and lower concentrations of cumate do not increase solubility, expressions were carried out with 50 μM cumate while altering other variables. Two expressions were performed at 37 ˚C and induced at high (OD600 = 0.8) and low (OD600 = 0.3) cell density (Figure 3.6). In addition, expressions were performed at three lower temperatures: 16 ˚C, room temperature, and 30 ˚C (Figures 3.7-3.9).



1 = FAD1 (soluble), 2 = GMC2 (soluble), 3 = RED1 (soluble), 4 = *E. coli* Fre (soluble), 5 = empty pCuminBB-cHis (soluble), 6 = FAD1 (total), 7 = GMC2 (total), 8 = RED1 (total), 9 = *E. coli* Fre (total), 10 = empty pCuminBB-cHis (total)

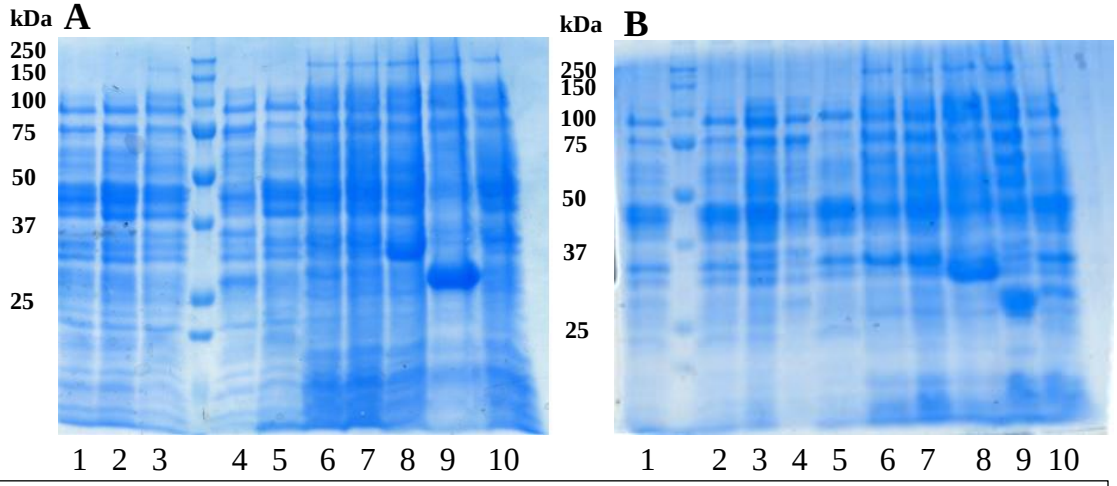
Figure 3.6 - SDS-PAGE analysis of flavin-binding enzyme expression at 37 ˚C. A) Induced at OD600 = 0.8 B) induced at OD600 = 0.3 See Table 3.3 for expected molecular weights of flavin-binding enzymes. Whether the protein fraction is soluble or total is indicated in parenthesis.

Altering the induction point does not appear to produce any soluble FAD1, GMC2, or *E. coli* Fre. However, a faint band is present in both of the RED1 soluble fractions at approximately the expected molecular weight of 32.6 kDa. It should be noted that the soluble fractions for RED1 have a number of bands at various molecular weights that are not normally seen or seen so darkly, thus this is not conclusive proof of RED1 being expressed in soluble form.



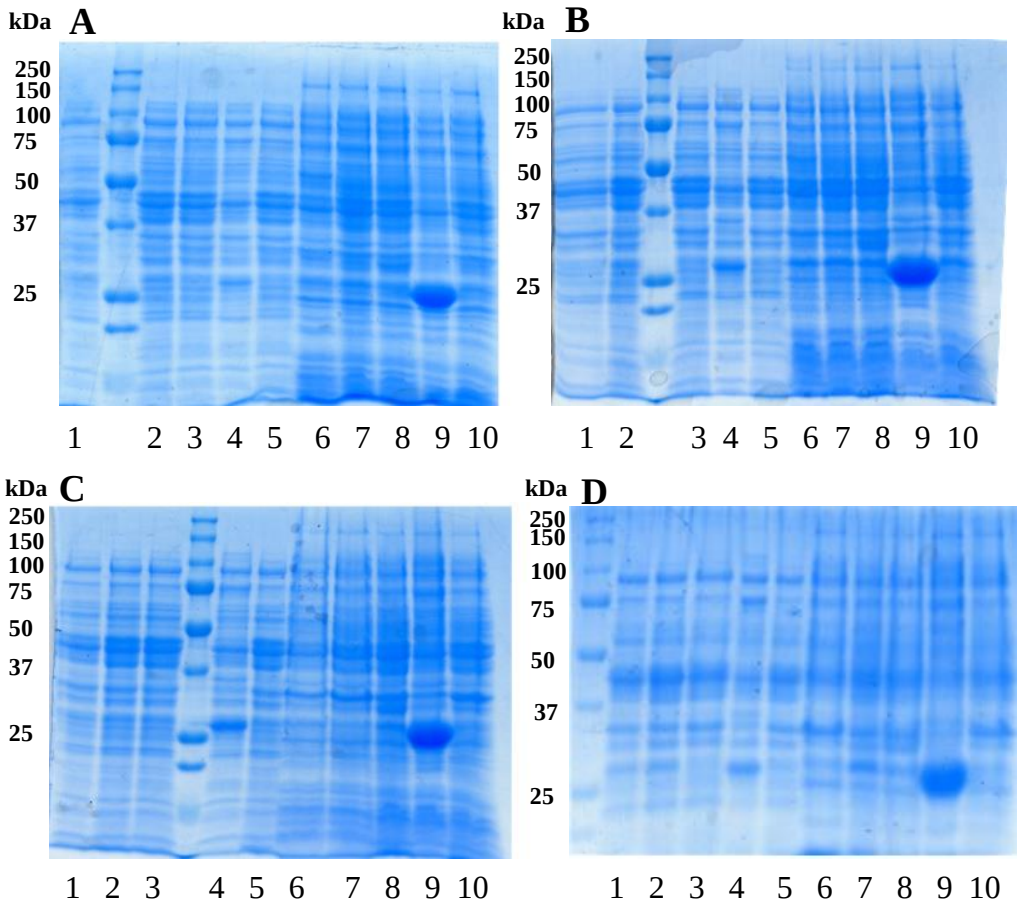
1 = FAD1 (soluble), 2 = GMC2 (soluble), 3 = RED1 (soluble), 4 = *E. coli* Fre (soluble), 5 = empty pCuminBB-cHis (soluble), 6 = FAD1 (total), 7 = GMC2 (total), 8 = RED1 (total), 9 = *E. coli* Fre (total), 10 = empty pCuminBB-cHis (total)

Figure 3.7 - SDS-PAGE analysis of flavin-binding enzyme expression at 50 µM cumate and 30˚C. A) Harvested 16 hours post induction (overnight) and B) 32 hours post induction. See Table 3.3 for expected molecular weights of flavin-binding enzymes. Whether the protein fraction is soluble or total is indicated in parenthesis.



1 = FAD1 (soluble), 2 = GMC2 (soluble), 3 = RED1 (soluble), 4 = *E. coli* Fre (soluble), 5 = empty pCuminBB-cHis (soluble), 6 = FAD1 (total), 7 = GMC2 (total), 8 = RED1 (total), 9 = *E. coli* Fre (total), 10 = empty pCuminBB-cHis (total)

Figure 3.8 - SDS-PAGE analysis of flavin-binding enzyme expression at 50 µM cumate and room temperature. A) Harvested 16 hours post induction and B) 32 hours post induction. See Table 3.3 for expected molecular weights of flavin-binding enzymes. Whether the protein fraction is soluble or total is indicated in parenthesis.



1 = FAD1 (soluble), 2 = GMC2 (soluble), 3 = RED1 (soluble), 4 = *E. coli* Fre (soluble), 5 = empty pCuminBB-cHis (soluble), 6 = FAD1 (total), 7 = GMC2 (total), 8 = RED1 (total), 9 = *E. coli* Fre (total), 10 = empty pCuminBB-cHis (total)

Figure 3.9 - SDS-PAGE analysis of flavin-binding enzyme expression at 50 µM cumate and 16 ˚C. A) Harvested 24 hours post induction, B) 48 hours post induction, C) 60 hours post induction and D) 72 hours post induction. See Table 3.3 for expected molecular weights of flavin-binding enzymes. Whether the protein fraction is soluble or total is indicated in parenthesis.

For both time points of the 30 ˚C and room temperature expressions, only RED1 and *E.coli* Fre appear to be expressed in the total protein fraction, and none of the flavin-binding enzymes are expressed in soluble form. At 16˚C, only *E. coli* Fre is expressed. However, it does appear to be expressed in the soluble fraction, with soluble expression reaching a significant level at 48 hours, maximum level at 60 hours, and decreasing by 72 hours. *E. coli* Fre appears to be expressed quite well in the total protein fraction, despite 50 μM cumate being used, which is somewhat in contrast to the results reported in Figure 3.5 (in which Fre expression decreased with increasing cumate concentration).

Co-expression of flavin-binding enzymes with chaperone proteins

Molecular chaperones are proteins which assist in the correct folding of other proteins. In *E. coli,* GroEL, GroES, DnaK, DnaJ, GrpE and trigger factor (Tf) are important chaperone proteins, GroEL working in conjunction with GroES (GroEL-GroES system) and DnaK, DnaJ, and GrpE forming another team (DnaK-DnaJ-GrpE system). When co-expressed with heterologous proteins, these chaperones have been shown to aid in the folding of recombinant proteins and in a number of cases cause them to be expressed in soluble form.[56,57,58] Takara® has produced a set of plasmids encoding different combinations of these chaperones (see Table 3.4 below), and these were co-expressed with the flavin-binding enzymes FAD1, GMC2, and RED1. Expression cultures were analyzed by SDS-PAGE. Note: Expressions of chaperone proteins alone established that the proteins from plasmid C3 (Table 3.5) did not express. Also, as RED1 appeared to possibly express in the soluble fraction of both high and low induction expressions (without chaperones) (see previous section), this expression was repeated.

Table 3.4 - Molecular weights of Takara® chaperone proteins.[55,56]

Chaperone Protein	GroEL	GroES	DnaK	DnaJ	GrpE	Tf
Approximate Molecular Weight	60 kDa	10 kDa	70 kDa	40 kDa	22 kDa (may appear around 29 kDa)	56 kDa

Table 3.5 - Takara® chaperone protein expression plasmids.[55,56]

Plasmid #	C1	C2	C3	C4	C5
Chaperone proteins expressed on plasmid	DnaK-DnaJ-GrpE GroES-GroEL	GroES-GroEL	DnaK-DnaJ-GrpE	GroES-GroEL-Tf	Tf
Expected SDS-PAGE Bands	70, 60, 40, 22/29, and 10 kDa	60 and 10 kDa	70, 40, and 22/29 kDa	60, 56, and 10 kDa	56 kDa

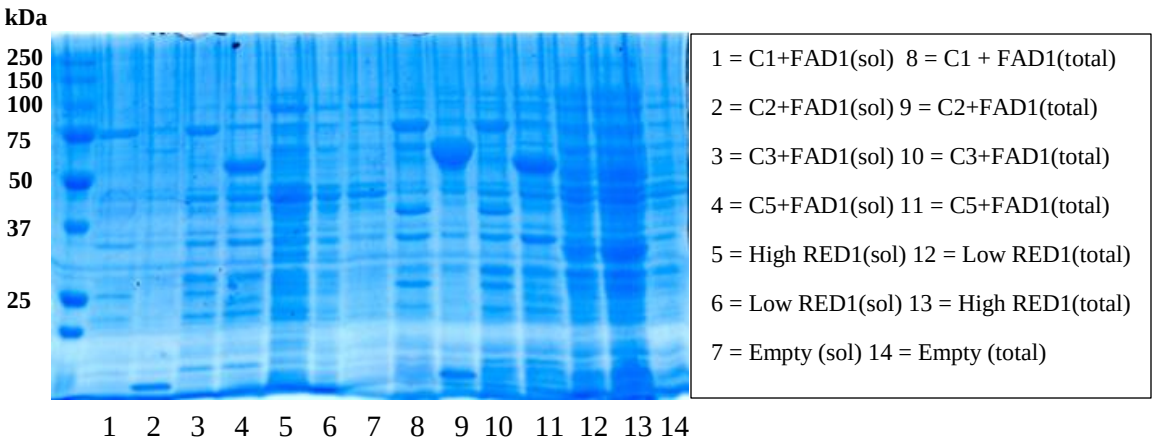


Figure 3.10 - SDS-PAGE analysis of FAD1 expression and chaperone co-expression at 16 hours post induction. See Tables 3.4 and 3.5 for expected molecular weights of chaperone proteins and chaperone plasmid descriptions. FAD1 Expected MW = 52.2 kDa. Note: total = total protein, sol = soluble protein.

After 16 hours, cultures were analyzed to determine if expression was occurring, with FAD1 and chaperone co-expressions analyzed (Figure 3.10). By this timepoint, it appeared that FAD1 was not expressed (no band observed at FAD1's expected molecular weight) and expressions were continued, with cultures analyzed again at 24 (Figure 3.11) and 48 hours (Figure 3.12) post induction.

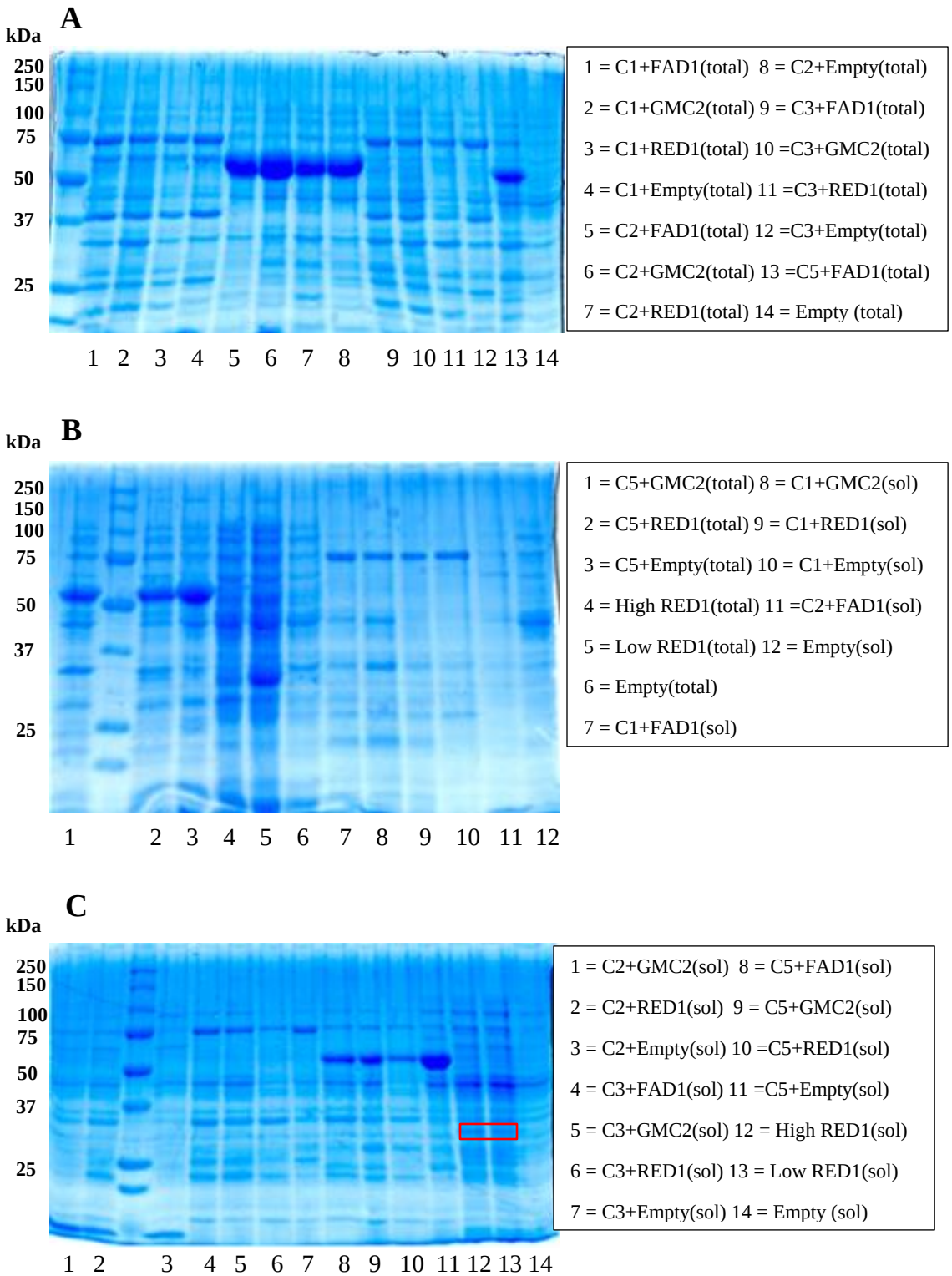


Figure 3.11 - SDS-PAGE analysis of flavin-binding enzyme and chaperone co-expression at 24 hours post induction (PI). See Tables 3.3, 3.4, and 3.5 for expected molecular weights and chaperone plasmid descriptions. Note: total = total protein, sol = soluble protein. Red box indicates potential RED1 soluble expression.

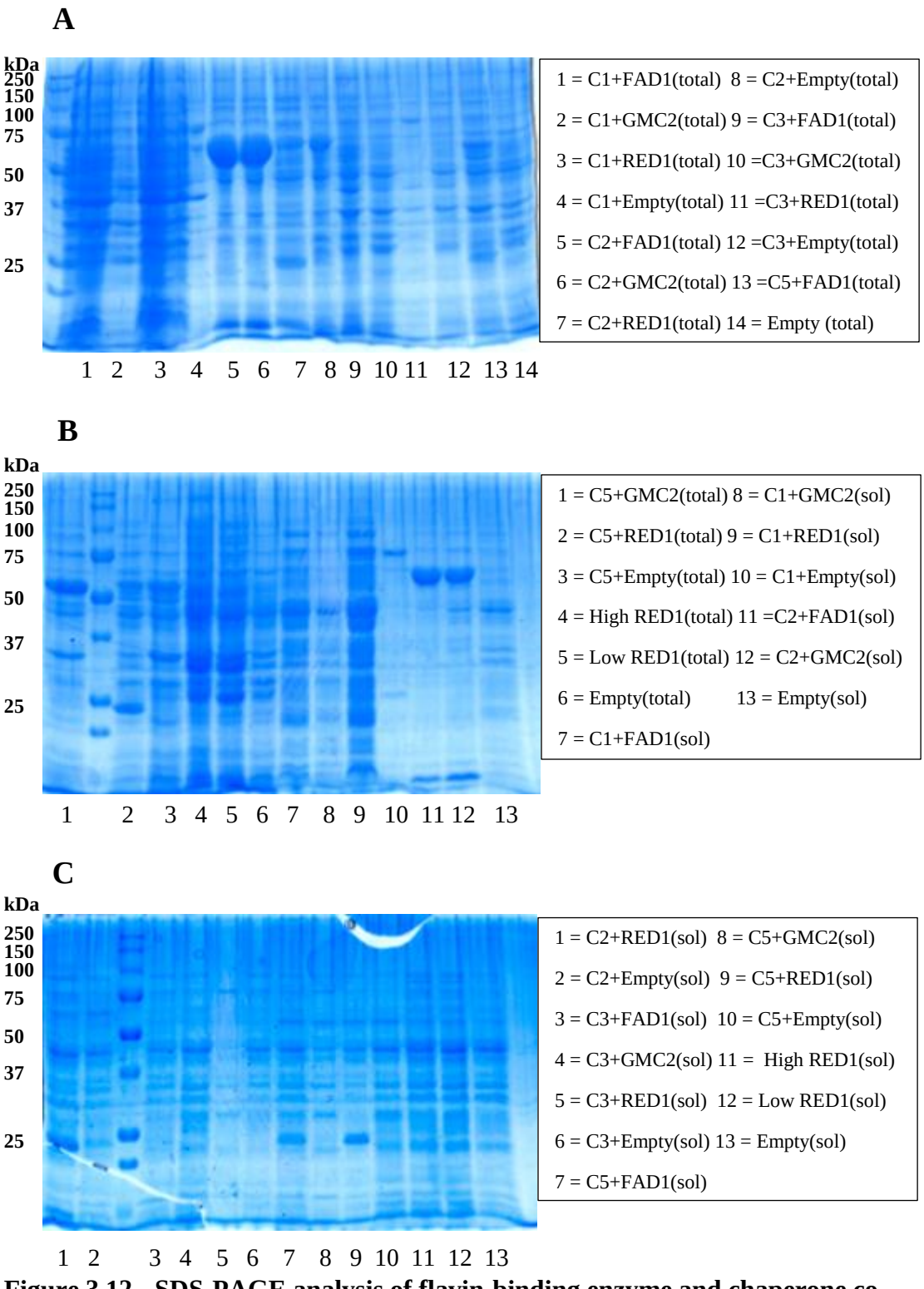


Figure 3.12 - SDS-PAGE analysis of flavin-binding enzyme and chaperone co-expression at 48 hours PI. See Tables 3.3, 3.4, and 3.5 for expected molecular weights and chaperone plasmid descriptions. Total=total protein, sol = soluble protein.

By 16 hours post induction, chaperone proteins appear to be expressed, however, there is no indication that flavin-binding enzymes are expressed even at 48 hours post induction. Thus, it would appear that chaperone expression inhibits flavin-binding enzyme co-expression. It should be noted that in lanes 11 and 12 of Figure 3.12 B (48 hours post induction), soluble bands appear that do not appear in other GroES-GroEL expressions. However, given the molecular weight of the bands (clearly higher than 50 kDa, probably approximately 60 kDa), this is likely soluble GroEL and not soluble FAD1 or GMC2. In the total protein fractions of the same expression (see lanes 5-8 of Figure 3.12 A), only the FAD1 and GMC2 GroEL-GroES coexpressions contain bands at the expected weight of GroEL. However, RED1 alone in high and low induction expressions (Expected MW = 32.6 kDa) does appear to be expressed by 24 hours (Figure 3.11 B lanes 4 and 5). The presence of a faint band (not present in the negative control, marked by a red box on gel) at the expected molecular weight in the soluble fractions (Figure 3.11 C lanes 12-14) indicates RED1 may be expressed in soluble form under these conditions, and thus purification could be attempted.

Conclusion

Initial expressions of flavin-binding enzymes from the Stehi7 gene cluster under the control of the cumate promoter resulted in no expression for FAD2, little to no expression for Omp7a, insoluble inclusion bodies for FAD1, GMC2, RED1, and *E. coli* Fre, and no soluble protein of any kind. In order to cause soluble expression, inductant (cumate) level, expression temperature, induction time, and expression time were varied. Varying the level of cumate used caused the total protein expression to vary in a roughly

proportional manner for the flavin-binding enzymes FAD1, GMC2, and RED1, but did not increase their solubility.

Interestingly, *E. coli* Fre appeared to be strongly expressed in all cultures (both those containing and not containing an *E. coli* Fre cumate inducible plasmid) without cumate, with expression decreasing at higher cumate levels. However, subsequent cultures with high levels (50 μM) of cumate still expressed Fre strongly. When cultures were grown at 16 ˚C, Fre expression at 50 μM cumate resulted in significant soluble Fre presence after 48 hours. Performing expressions at high and low induction points appears to potentially result in low levels of soluble RED1 protein, thus purification of RED1 and Fre can be attempted. Coexpression of FAD1, GMC2, and RED1 with the chaperone proteins GroEL, GroES, DnaJ, DnaK, GrpE, and Tf was also attempted in order to assist with folding and obtain soluble protein. However, it appears that coexpression with these chaperones inhibits flavin-binding enzyme expression.

Chapter 4: Conclusions and Future Directions

In pursuit of achieving *in vitro* biocatalysis of terpenoid compounds, an *Agrocybe aegerita* unspecific peroxygenase mutant (PaDa I-UPO) and flavin-binding enzymes from Stehi7 (*Stereum Hirsutum*) and Omp7 (*Omphalatus olearius*) Δ6-protoilludene synthase gene clusters were expressed in *S. cerevisiae* and *E. coli* respectively. The *S. cerevisiae* culture expressing PaDa I – UPO had the expected peroxidase activity against ABTS and peroxygenase activity against NBD, but no distinct protein band was observed at the expected molecular weight of 51.1 kDa vis à vis the empty vector control. Subsequent expressions at different temperatures (25 and 20 ˚C), different induction points (two, two and a half, and three doubling times), and six Kozak sequences (based on either the vertebrate or native *S. cerevisiae* Kozak sequence) yielded an increase of total activity to approximately 1.5 U/mL, with the highest activity of any single culture being 1.63 ± 0.058 U/mL or 47.9% of the reported total activity (3.4 U/mL) achieved in the original paper.[19]

Precipitation of the PaDa I-UPO culture supernatant with 30% ammonium sulfate appeared to remove many of the contaminating proteins without reducing total activity, however fractional precipitation up to 60% ammonium sulfate concentrated but did not further purify PaDa I – UPO, with the PaDa I – UPO and empty vector control fractions containing the same visible protein bands by SDS-PAGE. Ion exchange chromatography resulted in no further purification, as PaDa I – UPO was not retained by the cation exchange column. Though it was not purified further, PaDa I – UPO was demonstrated to have activity against terpenes by GC/MS analysis, converting limonene to limonene

epoxide and carveol (replicating the findings of Sebastian, et al)[23] and Δ6-protoilludene to a probably oxygenated compound with a molecular weight of 220 daltons.

In order to increase production of PaDa I – UPO and obtain a distinct enzyme band by SDS – PAGE, a logical next step will be to express it in the methylotrophic yeast *Pichia pastoris.* PaDa I – UPO has been expressed in *P. pastoris* in the literature, and while this did not increase PaDa I – UPO production in flask fermentation it did result in a 27 fold increase over *S. cerevisiae* in fed batch fermentation.[29] It is also possible that a distinct band for PaDa I – UPO is present even without increasing production, but this band is hidden by a contaminant protein band at the same molecular weight in both the empty vector control and the expression culture. As PaDa I – UPO is glycosylated (Molina – Espeja et al., 2014) its molecular weight might vary from the expected value of 51.1 kDa, and thus it could be present in the 75 kDa band seen in expression cultures and empty vector controls. Mass spectrometry of this protein band would be able to ascertain whether it contains PaDa I – UPO. Ultimately PaDa I – UPO should be purified in order to certify that it alone is responsible for activity against Δ6-protoilludene or other terpenes, and to determine if it would be more active in purified form. As PaDa I – UPO has been purified using ion exchange chromatography[19] this will continue to be pursued, with additional adjustments made to buffers and the type of column used so that the enzyme remains on the column. Once PaDa I – UPO is produced in sufficient quantity and purified, the final goal is to use it as a scaffold modifying enzyme in a biocatalytic cascade that can convert Δ6-protoilludene and possibly other terpenes to bioactive terpenoids. For this to be feasible, PaDa I – UPO should remain active over a relatively long period of time and convert a significant amount of Δ6-protoilludene to the modified

product in a repeatable manner, neither of which it currently does. It has been demonstrated that PaDa I – UPO remains active for a longer period of time when H_2O_2 is added slowly into the reaction,[47] thus for future terpene modification reactions H_2O_2 should be added slowly through a device such as a pump.

As many terpenoids contain multiple oxygen atoms, possessing a toolbox of scaffold modifying enzymes with the ability to oxygenate different positions on a molecule is essential. As cytochrome P450s have proven difficult to express in active form, the flavin-binding enzymes FAD1, FAD2, GMC2, RED1, and Omp7a were investigated for potential use alongside *E. coli* Fre. While FAD1, GMC2, RED1, and Fre expressed strongly in *E. coli*, initial expressions were insoluble. Altering conditions such as inductant concentration, expression temperature, induction time, and expression time resulted in soluble expression of Fre and possible soluble expression of RED1. The next step will be to attempt purification of Fre and RED1 using affinity chromatography, likely with nickel ion resin. It is not uncommon for heterologous proteins to be in inclusion bodies when expressed in *E. coli,* and there remain a number of techniques that can be attempted in order to produce soluble FAD1 and GMC2. One avenue that should be attempted is adding a flavin source to the expression media, as this can aid in the expression of soluble flavoenzymes.[60] Other methods which can be tried include molecular techniques to lower expression rate (i.e. weaker promoter or using a lower copy number plasmid), performing expressions with various other media types besides Luria-broth, and denaturing and refolding the insoluble protein *in vitro* through the addition and slow removal of a denaturant. Another step that should be taken would be to co-express Δ6-protoilludene synthase with FAD1, GMC2, and RED1 and observe

whether these enzymes are active *in vivo* against Δ6-protoilludene in spite of their insolubility *in vitro,* with activity being measured by GC/MS analysis of the culture headspace. Ultimately, while I was not able to complete a biosynthetic pathway to Δ6-protoilludene derivatives or other sesquiterpenoids, there was some promise in PaDa I – UPO's activity against Δ6-protoilludene. Potentially, this enzyme could be used in a biosynthetic cascade. In terms of the flavin-binding enzymes, while no scaffold modifying enzymes were shown to be active *in vitro,* first steps were taken toward their heterologous expression and isolation.

Bibliography

1.Dias, D. A.; Urban, S.; Roessner, U., A historical overview of natural products in drug discovery. Metabolites 2012, 2 (2), 303-36.

2. Paddon, C. J.; Keasling, J. D., Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. Nature reviews. Microbiology 2014, 12 (5), 355-67.

3.Lobanovska, M.; Pilla, G., Penicillin's Discovery and Antibiotic Resistance: Lessons for the Future? The Yale journal of biology and medicine 2017, 90 (1), 135-145.

4.Luo, Y.; Cobb, R. E.; Zhao, H., Recent advances in natural product discovery. Current opinion in biotechnology 2014, 30, 230-7.

5.Schmidt-Dannert, C., Biosynthesis of Terpenoid Natural Products in Fungi. Adv Biochem Eng Biotechnol 2015, 148: 19–61.

6. Ingy, A.l Quax, W. J., A Glimpse into the Biosynthesis of Terpenoids. International Conference on Natural Resources and Life Sciences. 2017

7. Engels, B.; Heinig, U.; Grothe, T.; Stadler, M.; Jennewein, S., Cloning and characterization of an Armillaria gallica cDNA encoding protoilludene synthase, which catalyzes the first committed step in the synthesis of antimicrobial melleolides. The Journal of biological chemistry 2011, 286 (9), 6871-8.

8. Mora-Pale, M.; Sanchez-Rodriguez, S. P.; Linhardt, R. J.; Dordick, J. S.; Koffas, M. A., Metabolic engineering and in vitro biosynthesis of phytochemicals and non-natural analogues. Plant science : an international journal of experimental plant biology 2013, 210, 10-24.

9. Zhou, X.; Zhu, H.; Liu, L.; Lin, J.; Tang, K., A review: recent advances and future prospects of taxol-producing endophytic fungi. Applied microbiology and biotechnology 2010, 86 (6), 1707-17.

10. Howat, S.; Park, B.; Oh, I. S.; Jin, Y. W.; Lee, E. K.; Loake, G. J., Paclitaxel: biosynthesis, production and future prospects. New biotechnology 2014, 31 (3), 242-5.

11. Ma, B.J.; Wu, T.T.; Ruan, Y.; Shen, J.W.; Zhou, H.; Yu, H.; Zhao, X., Antibacterial and antifungal activity of scabronine G and H in vitro. Mycology 2010, 1(3), 200-203.

12. Liermann, J. C.; Thines, E.; Opatz, T.; Anke, H., Drimane sesquiterpenoids from Marasmius sp. inhibiting the conidial germination of plant-pathogenic fungi. Journal of natural products 2012, 75 (11), 1983-6.

13. Quin, M. B.; Flynn, C. M.; Schmidt-Dannert, C., Traversing the fungal terpenome. Natural product reports 2014, 31 (10), 1449-73.

14. Ding, Y. X.; Ou-Yang, X.; Shang, C. H.; Ren, A.; Shi, L.; Li, Y. X.; Zhao, M. W., Molecular cloning, characterization, and differential expression of a farnesyl-diphosphate synthase gene from the basidiomycetous fungus Ganoderma lucidum. Bioscience, biotechnology, and biochemistry 2008, 72 (6), 1571-9.

15. Agger, S.; Lopez-Gallego, F.; Schmidt-Dannert, C., Diversity of sesquiterpene synthases in the basidiomycete Coprinus cinereus. Molecular microbiology 2009, 72 (5), 1181-95.

16. Wawrzyn, G. T.; Quin, M. B.; Choudhary, S.; Lopez-Gallego, F.; Schmidt-Dannert, C., Draft genome of *Omphalotus olearius* provides a predictive framework for sesquiterpenoid natural product biosynthesis in Basidiomycota. Chemistry & biology 2012, 19 (6), 772-83.

17. Quin, M. B.; Flynn, C. M.; Wawrzyn, G. T.; Choudhary, S.; Schmidt-Dannert, C., Mushroom hunting by using bioinformatics: application of a predictive framework facilitates the selective identification of sesquiterpene synthases in Basidiomycota. Chembiochem : a European journal of chemical biology 2013, 14 (18), 2480-91.

18. Hausjell, J.; Halbwirth, H.; Spadiut, O., Recombinant production of eukaryotic cytochrome P450s in microbial cell factories. Bioscience reports 2018, 38 (2).

19. Molina-Espeja, P.; Garcia-Ruiz, E.; Gonzalez-Perez, D.; Ullrich, R.; Hofrichter, M.; Alcalde, M., Directed evolution of unspecific peroxygenase from *Agrocybe aegerita*. Applied and environmental microbiology 2014, 80 (11), 3496-507.

20. Hofrichter, M.; Kellner, H.; Pecyna, M. J.; Ullrich, R., Fungal unspecific peroxygenases: heme-thiolate proteins that combine peroxidase and cytochrome P450 properties. Advances in experimental medicine and biology 2015, 851, 341-68.

21. Ullrich, R.; Nuske, J.; Scheibner, K.; Spantzel, J.; Hofrichter, M., Novel haloperoxidase from the agaric basidiomycete *Agrocybe aegerita* oxidizes aryl alcohols and aldehydes. Applied and environmental microbiology 2004, 70 (8), 4575-81.

22. Peter, S.; Kinne, M.; Wang, X.; Ullrich, R.; Kayser, G.; Groves, J. T.; Hofrichter, M., Selective hydroxylation of alkanes by an extracellular fungal peroxygenase. The FEBS journal 2011, 278 (19), 3667-75.

23.Peter, S.; Kinne, M.; Ullrich, R.; Kayser, G.; Hofrichter, M., Epoxidation of linear, branched and cyclic alkenes catalyzed by unspecific peroxygenase. Enzyme and microbial technology 2013, 52 (6-7), 370-6.

24. Ullrich, R.; Hofrichter, M., The haloperoxidase of the agaric fungus *Agrocybe aegerita* hydroxylates toluene and naphthalene. FEBS letters 2005, 579 (27), 6247-50.

25.Karich, A.; Kluge, M.; Ullrich, R.; Hofrichter, M., Benzene oxygenation and oxidation by the peroxygenase of *Agrocybe aegerita*. AMB Express 2013, 3 (1), 5.

26. Aranda, E.; Ullrich, R.; Hofrichter, M., Conversion of polycyclic aromatic hydrocarbons, methyl naphthalenes and dibenzofuran by two fungal peroxygenases. Biodegradation 2010, 21 (2), 267-81.

27. Kinne, M.; Poraj-Kobielska, M.; Ralph, S. A.; Ullrich, R.; Hofrichter, M.; Hammel, K. E., Oxidative cleavage of diverse ethers by an extracellular fungal peroxygenase. The Journal of biological chemistry 2009, 284 (43), 29343-9.

28. Wang, Y.; Lan, D.; Durrani, R.; Hollmann, F., Peroxygenases en route to becoming dream catalysts. What are the opportunities and challenges? Current opinion in chemical biology 2017, 37, 1-9.

29. Molina-Espeja, P.; Ma, S.; Mate, D. M.; Ludwig, R.; Alcalde, M., Tandem-yeast expression system for engineering and producing unspecific peroxygenase. Enzyme and microbial technology 2015, 73-74, 29-33.

30. Kluge, M.; Ullrich, R.; Scheibner, K.; Hofrichter, M.; Stereoselective benzylic hydroxylation of alkylbenzenes and epoxidation of styrene derivatives catalyzed by the peroxygenase of *Agrocybe aegerita*. Green Chemistry 2012, 14(2), 440-446.

31. Hofrichter, M.; Ullrich, R.; Pecyna, M. J.; Liers, C.; Lundell, T., New and classic families of secreted fungal heme peroxidases. Applied microbiology and biotechnology 2010, 87 (3), 871-97.

32. Conesa, A.; van De Velde, F.; van Rantwijk, F.; Sheldon, R. A.; van Den Hondel, C. A.; Punt, P. J., Expression of the Caldariomyces fumago chloroperoxidase in Aspergillus niger and characterization of the recombinant enzyme. The Journal of biological chemistry 2001, 276 (21), 17635-40.

33. Anh, D. H.; Ullrich, R.; Benndorf, D.; Svatos, A.; Muck, A.; Hofrichter, M., The coprophilous mushroom Coprinus radians secretes a haloperoxidase that catalyzes aromatic peroxygenation. Applied and environmental microbiology 2007, 73 (17), 5477-85.

34. Grobe, G.; Ullrich, R.; Pecyna, M. J.; Kapturska, D.; Friedrich, S.; Hofrichter, M.; Scheibner, K., High-yield production of aromatic peroxygenase by the agaric fungus Marasmius rotula. AMB Express 2011, 1 (1), 31.

35. Wang, X.; Peter, S.; Kinne, M.; Hofrichter, M.; Groves, J. T., Detection and kinetic characterization of a highly reactive heme-thiolate peroxygenase compound I. Journal of the American Chemical Society 2012, 134 (31), 12897-900.

36. Wang, X.; Peter, S.; Ullrich, R.; Hofrichter, M.; Groves, J. T., Driving force for oxygen-atom transfer by heme-thiolate enzymes. Angewandte Chemie (International ed. in English) 2013, 52 (35), 9238-41.

37. Poraj-Kobielska, M.; Kinne, M.; Ullrich, R.; Scheibner, K.; Hofrichter, M., A spectrophotometric assay for the detection of fungal peroxygenases. Analytical biochemistry 2012, 421 (1), 327-9.

38. Cai, B.; Liu, X.; Zou, J.; Xiao, J.; Yuan, B.; Li, Fei.; Cheng, Q., Multi-wavelength spectrophotometric determination of hydrogen peroxide in water with peroxidase-catalyzed oxidation of ABTS. Chemosphere 2018, 193, 833-839.

39. Partow, S.; Siewers, V.; Bjorn, S.; Nielsen, J.; Maury, J., Characterization of different promoters for designing a new expression vector in Saccharomyces cerevisiae. Yeast (Chichester, England) 2010, 27 (11), 955-64.

40. Schneider, J. C.; Guarente, L., Vectors for expression of cloned genes in yeast: regulation, overproduction, and underproduction. Methods in enzymology 1991, 194, 373-88.

41. Wingfield, P. T., Overview of the purification of recombinant proteins. Current protocols in protein science 2015, 80, 6.1.1-35.

42. Graslund, S.; Nordlund, P.; Weigelt, J.; Hallberg, B. M.; Bray, J.; Gileadi, O.; Knapp, S.; Oppermann, U.; Arrowsmith, C.; Hui, R.; Ming, J.; dhe-Paganon, S.; Park, H. W.; Savchenko, A.; Yee, A.; Edwards, A.; Vincentelli, R.; Cambillau, C.; Kim, R.; Kim, S. H.; Rao, Z.; Shi, Y.; Terwilliger, T. C.; Kim, C. Y.; Hung, L. W.; Waldo, G. S.; Peleg, Y.; Albeck, S.; Unger, T.; Dym, O.; Prilusky, J.; Sussman, J. L.; Stevens, R. C.; Lesley, S. A.; Wilson, I. A.; Joachimiak, A.; Collart, F.; Dementieva, I.; Donnelly, M. I.; Eschenfeldt, W. H.; Kim, Y.; Stols, L.; Wu, R.; Zhou, M.; Burley, S. K.; Emtage, J. S.; Sauder, J. M.; Thompson, D.; Bain, K.; Luz, J.; Gheyi, T.; Zhang, F.; Atwell, S.; Almo, S. C.; Bonanno, J. B.; Fiser, A.; Swaminathan, S.; Studier, F. W.; Chance, M. R.; Sali, A.; Acton, T. B.; Xiao, R.; Zhao, L.; Ma, L. C.; Hunt, J. F.; Tong, L.; Cunningham, K.; Inouye, M.; Anderson, S.; Janjua, H.; Shastry, R.; Ho, C. K.; Wang, D.; Wang, H.; Jiang, M.; Montelione, G. T.; Stuart, D. I.; Owens, R. J.; Daenke, S.; Schutz, A.; Heinemann, U.; Yokoyama, S.; Bussow, K.; Gunsalus, K. C., Protein production and purification. Nature methods 2008, 5 (2), 135-46.

43.Kozak, M., Point mutations close to the AUG initiator codon affect the efficiency of translation of rat preproinsulin in vivo. Nature 1984, 308 (5956), 241-6.

44. Kozak, M., An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic acids research 1987, 15 (20), 8125-48.

45.Hamilton, R.; Watanabe, C. K.; de Boer, H. A., Compilation and comparison of the sequence context around the AUG startcodons in Saccharomyces cerevisiae mRNAs. Nucleic acids research 1987, 15 (8), 3581-93.

46. National Institute of Advanced Industrial Science and Technology (AIST). Spectral Database for Organic Compounds (SDBS). SDBSWeb : http://sdbs.db.aist.go.jp

47. Karich, A.; Scheibner, K.; Ulrich, R.; Hofrichter, M., Exploring the catalase activity of unspecific peroxygenases and the mechanism of peroxide-dependent heme destruction. Journal of Molecular Catalysis B: Enzymatic 2016, 134, 236-246.

48. Huijbers, M. M.; Montersino, S.; Westphal, A. H.; Tischler, D.; van Berkel, W. J., Flavin dependent monooxygenases. Archives of biochemistry and biophysics 2014, 544, 2-17.

49. Hefti, M. H.; Vervoort, J.; van Berkel, W. J., Deflavination and reconstitution of flavoproteins. European journal of biochemistry 2003, 270 (21), 4227-42.

50. Iida, K.; Cox-Foster, D. L.; Yang, X.; Ko, W. Y.; Cavener, D. R., Expansion and evolution of insect GMC oxidoreductases. BMC evolutionary biology 2007, 7, 75.

51. Keller, N. P., Translating biosynthetic gene clusters into fungal armor and weaponry. Nature chemical biology 2015, 11 (9), 671-7

52. Massey, V., Activation of molecular oxygen by flavins and flavoproteins. The Journal of biological chemistry 1994, 269 (36), 22459-62.

53. Huang, C. H.; Lai, W. L.; Lee, M. H.; Chen, C. J.; Vasella, A.; Tsai, Y. C.; Liaw, S. H., Crystal structure of glucooligosaccharide oxidase from Acremonium strictum: a novel flavinylation of 6-S-cysteinyl, 8alpha-N1-histidyl FAD. The Journal of biological chemistry 2005, 280 (46), 38831-8.

54. Spyrou, G.; Haggard, L.E.; Krook, M., Jornvall, H., Nilsson, E., Reichard, P.,Characterization of the flavin reductase gene (fre) of Escherichia coli and construction of a plasmid for overproduction of the enzyme. The Journal of Bacteriology 1991, 173(12), 3673

55. Mullick, A.; Xu, Y.; Warren, R.; Koutroumanis, M.; Guilbault, C.; Broussau, S.; Malenfant, F.; Bourget, L.; Lamoureux, L.; Lo, R.; Caron, A.; Pilotte, A.; Massie, B., The cumate gene-switch: a system for regulated expression in mammalian cells. BMC Biotechnology 2006, 6:43.

56. Nishihara, K.; Kanemori, M.; Kitagawa, M.; Yanagi, H.; Yura, T., Chaperone coexpression plasmids: differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in Escherichia coli. Applied and environmental microbiology 1998, 64 (5), 1694-9.

57.Nishihara, K.; Kanemori, M.; Yanagi, H.; Yura, T., Overexpression of trigger factor prevents aggregation of recombinant proteins in Escherichia coli. Applied and environmental microbiology 2000, 66 (3), 884-9.

58. Thomas, J. G.; Ayling, A.; Baneyx, F., Molecular chaperones, folding catalysts, and the recovery of active recombinant proteins from E. coli. To fold or to refold. Applied biochemistry and biotechnology 1997, 66 (3), 197-238.

59. Durairaj, P.; Hur, J. S.; Yun, H., Versatile biocatalysis of fungal cytochrome P450 monooxygenases. Microbial cell factories 2016, 15 (1), 125.

60. Brunelle, A.; Bi, Y.; Lin, J.; Russell, B.; Luy, L.; Berkman, C.; Cashman, J., Characterization of Two Human Flavin-Containing Monooxygenase (Form 3) Enzymes Expressed in Escherichia coli as Maltose Binding Protein Fusions. Drug Metabolism and Disposition 1997, 25 (8) 1001-1007.

Supplemental Materials

PaDa I – UPO assay against valencene

In addition to assaying PaDa I – UPO against limonene and Δ6-protoilludene, PaDa I – UPO was also assayed against the sesquiterpene valencene. The reaction set up and GC/MS trial were identical to that described for limonene in the materials and methods section of Chapter 1.

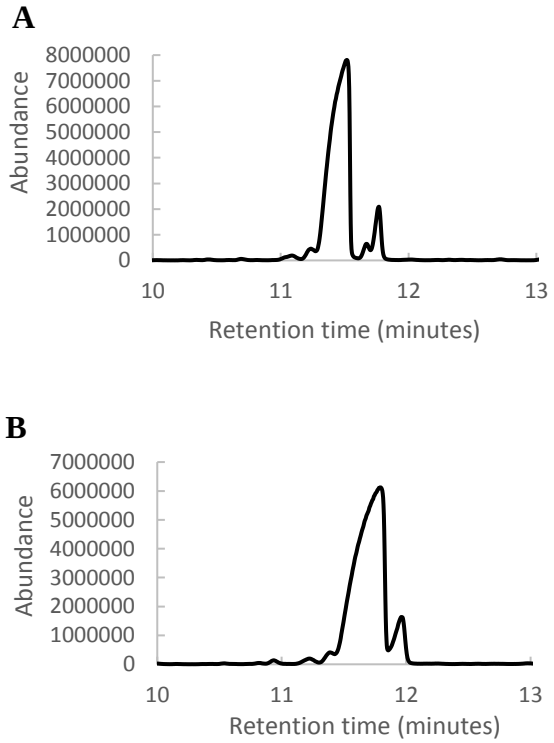


Figure S1. **A) Empty vector control + valencene, B) PaDa I – UPO + valencene.**

Both chromatograms appeared to show multiple peaks, but all peaks in both chromatograms were identified as valencene by NIST. It would thus appear that PaDa I – UPO lacks activity against valencene.

Yeast transformation protocol

The protocol used to transform pESC-ura plasmids into *S. cerevisiae* (Chapter 2) is described here. Note this protocol was developed by Dr. Sarah Perdue of the Claudia Schmidt-Dannert lab.

1. Scrape several fresh yeast colonies (approximately 25 µL of inoculum) from a plate into 1 mL of sterile H_2O.
2. Pellet the cells for 5 seconds
3. Resuspend the cells in 1 mL 100 mM LiAc and incubate at room temperature for 5 minutes. For each transformation you are performing, boil 10 µL of 10 mg/mL salmon sperm carrier DNA for 5 minutes.
4. For each transformation, remove 250 µL of the resuspended yeast cells. Pellet this aliquot for five seconds, and remove supernatant.
5. Add the following sterile components to each cell pellet being transformed:
    - 240 µL of 50% PEG 3350
    - 36 µL of 1.0 M LiAc
    - 10 µL of 10 mg/mL salmon sperm carrier DNA (boiled)
    - 64 µL of water
6. Add DNA. For an intact plasmid, add 2.5 µL of a standard vector miniprep.
7. Vortex reaction mixture for 1 minute to resuspend the pellet
8. Incubate at 42 ˚C for 20 minutes.
9. Pellet cells for 10 seconds at top speed. Discard the supernatant.
10. Resuspend pellet in 400 µL H_2O.
11. Plate 5 – 200 µL of the mixture from step 10 on selective media. Incubate at 30 ˚C for 2-3 days.